Universidad
Carlos III de Madrid
www.uc3m.es

TESIS DOCTORAL

Three Essays on Applied Economics

Autor:

Yunrong Li

Director:

Ricardo Mora

DEPARTAMENTO DE ECONOMIA

Getafe, julio de 2014

Universidad
Carlos III de Madrid
www.uc3m.es

TESIS DOCTORAL

Three Essays on Applied Economics

Autor: *Yunrong Li*

Director: Ricardo Mora

Firma del Tribunal Calificador:

Firma

Presidente:

Vocal:

Secretario:

Calificación:

Getafe, 15 de julio de 2014

Acknowledgements

First and foremost, I would like to thank my advisor, professor Ricardo Mora. It has been an honor to be his Ph.D. student. He is one of the smartest and funniest people I know. I appreciate all his contributions of time, insightful discussions and supports, both academic and economic, during the past four years of my Ph.D. experience. The joy and enthusiasm he has for the research was contagious. His rigorous attitude has deeply infected and inspired me.

I also want to thank professor Javier Ruiz-Castillo. I am very much impressed by his productivity and efficiency in the research work. During the two years of being a research assistant for him, I have not only developed my skills of processing data sets, but also benefited from his inspirational ideas and persistent pursuit of the truth.

Many thanks to Natalia Fabra, Antoine Leoper and Juanjo Dolado, for all the instrumental comments on my job market paper. Special thanks to the members of the thesis committee who devote a lot to improve my thesis. Talks to Heiko, Ming-jin, Robert, Efi and Román are very much appreciated.

I wish to thank all the administrative staff at the department. Thanks a lot to Arancha, Angelica and Sonia for your collaborations and kind help, especially during the job market period.

Thank you very much my Ph.D. fellows, Robert, Francesco, Anil, Lian, Joseph, Victor, etc., for all the kindness and the moments we share together. Special thanks to Rui, Xiaojun and Mian, thank you for being there when I had hard time. I sincerely cherish the joy we had together. Thank you Hang for the cake for my last birthday in Getafe. Thank you Tao for being a companion at the cafeteria. Thank you Ming-Jin, sharing the flat with you was really a memorable interval of time. Hope that we can keep in touch in the future.

感谢我的爸爸妈妈，你们的理解和支持，是我求学道路上最大的支撑，你们的坚强和善良，影响着我做人做事的准则。感谢我的姐姐和姐夫，在我离家的日子，陪伴和照顾父母，在我回家时，对我无微不至地关心，听我倾诉。在此，我把这篇论文献给你们，我最爱的家人。

Abstract

This thesis focuses on the application of empirical research methods to different economic topics. The first chapter examines production effects of subsidies with different characteristics. The second chapter evaluates the impact of an oldage pension program on the welfare of the recipient's family members. The third chapter applies an income inequality model to study the influence of differences in citation practices across scientific fields on the overall citation inequality.

Chapter 1, "Differential Effects on Output Levels of Binding and non-Binding Subsidies under Capitalization". Subsidies on outputs or inputs are usually production-promoting by lowering the marginal cost. However, if subsidies are binding, i.e. outputs or inputs are partially subsidized, subsidies don't affect the output level. If subsidies capitalize into input prices, i.e. subsidies benefit both the recipients and input providers, outputs will be negatively affected. My paper contributes by empirically assessing production effects of subsidies taking into account both bindingness and capitalization. I study cattle payments under the Common Agricultural Policy (CAP) implemented in the European Union (EU). I set up a simple model to analyze production effects of these payments. I also estimate the effects with Spanish farm-level data. CAP 1992 and Agenda 2000 are two policy programs of the CAP. Both are designed to reduce over-production in agriculture. Estimation results suggest that cattle payments have negative impacts on outputs when they are binding under CAP 1992, and positive impacts when they are non-binding under Agenda 2000.

Chapter 2, "Reassessing the Differential Impact of Grandmothers and Grandfathers: The Old Age Program in Nepal" (co-authored with Ricardo Mora). We study the effects on infant mortality of the introduction in 1995 of a non-contributory universal pension scheme in Nepal known as the Old age Allowance Program. We use cross-sectional data from the 1996 and 2001 Nepal Demographic and Health Surveys. Following a standard diff-in-diffs approach, we find positive and significant effects on survival rates for the presence in the same household of a female beneficiary while negative and sometimes significant effects for the presence of a male beneficiary. When we conduct pre-treatment common trend tests, we find that we cannot reject it for the case of the female beneficiaries but we strongly reject it for the case of male beneficiaries. Following Mora and Reggio (2012), we then propose a more flexible model and identification strategy and find that there are no differences in the female and the male beneficiary effects. We interpret these results as suggestive that cross-sectional analysis may bias downwards the estimates of the effect of grandfathers because of gender differences in endogenous household formation.

Chapter 3 is a combination of two closely related papers, namely "The Measurement of the Effect on Citation Inequality of Differences in Citation Practices across Scientific Fields" (co-authored with Juan A. Crespo and Javier Ruiz-Castillo, published in PLoS ONE 8(3): e58727 (2013)), and "The Effect on Citation Inequality

of Differences in Citation Practices at the Web of Science Subject Category Level" (co-authored with Juan A. Crespoa, Neus Herranz and Javier Ruiz-Castillo, published in Journal of the American Society for Information Science and Technology, 65:1244-1256, (June 2014)). We introduce a novel method for measuring which part of overall citation inequality can be attributed to differences in citation practices across scientific fields. In addition, we implement an empirical strategy for making meaningful comparisons between the numbers of citations received by articles in different scientific fields. Using a dataset of 4.4 million articles published in 1998-2003 with a five-year citation window, we find that differences in citation practices between the 22 fields account for about 14% of overall citation inequality. When the classification system goes from 22 fields to 219 sub-fields, the effect on citation inequality increases to about 18%. For comparisons of citation counts across fields, we provide a set of exchange rates (ERs) to express citations in any field into citations in the all-fields case. When the raw citation data are normalized with our ERs, the effect of differences in citation practices is reduced to around 2% of overall citation inequality in the case of 22 fields. In the case of 219 sub-fields with the fractional strategy, the normalization of the raw data using the ERs (or sub-field mean citations) as normalization factors reduces the effect to 3.8% (3.4%) of overall citation inequality. The results with the fractional strategy are essentially replicated when we adopt a multiplicative approach.

Chapter 1: Differential Effects on Output Levels of Binding and non-Binding Subsidies under Capitalization

1. Introduction

According to economic intuitions, subsidies on outputs or inputs are usually production-promoting by lowering the marginal cost. However, if subsidies are binding which means that outputs or inputs are partially subsidized, the marginal product (hence the total output) is determined by the standard equalization condition between price and marginal cost. Hence, binding subsidies have no effect on the output. Production effects of subsidies also depend on whether subsidies capitalize into input prices (i.e. subsidies benefit both the recipients and inputs providers). If subsidies induce input prices to rise, higher marginal cost will result in a negative impact on the output. If bindingness and capitalization happen at the same time, the output will decrease compared with the situation without subsidies. If subsidies are non-binding and capitalization is partial (i.e. subsidies don't go entirely to input providers), the output will be greater than that without subsidies.

Capitalization of subsidies rises from two facts. First, if subsidies promote the production, demands for inputs will increase and thus the input prices (Rolph (1952), Floyd (1965), Roberts (2003), Patton (2008), Kirwan (2009)). Roberts (2003) also points out that higher input prices will lead to a low supply respond. Second, subsidies increase the value of marginal product of inputs, where the value is the sum of the market price and the subsidy (Rolph (1952)). This can happen even if subsidies are binding. In agricultural studies, there are empirical evidences about the extent to which subsidies capitalize. Kirwan (2009) studies direct payments on crops in the US and concludes that landlords capture around 20% of the marginal subsidy dollar through higher rents, tenants' net returns account for about 70% of the subsidy, while the rest 10% may be extracted by other input providers. Patton (2008) studies direct payments on beef cattle in the European Union (EU) during 1994 to 2002 and finds that around 40% of two types of payments goes into land rents. These studies focus on the distribution of subsidies between the recipients and input providers, especially landowners, but not on production effects of subsidies.

There are just a few studies concerning bindingness of subsidies while analyzing production effects of subsidies. Chincinga (2008) studies a subsidy program that provides limited amount of fertiliser to farmers and finds a positive effect on maize yields. The positive effect may come from the fact that the subsidy is not binding at least for some farmers. In other words, the subsidized amount is enough for the usage. Gohin (2006) models production effects of cattle payments under CAP 1992 and Agenda 2000, both of which are reform programs of the European Common Agricultural Policy (CAP) implemented in the EU. He sets up two model specifications about farmers' production decisions distinguishing whether the payments are limited by farm-level quotas. His calibration results show that the output with binding payments is lower than that with non-binding payments, and a decrease in binding payments has no effect on the output. However, he doesn't consider possible capitalization of cattle payments that may negatively affect the output. This makes his model incomplete.

So far there is no literature taking into account both bindingness and capitalization when studying production effects of subsidies. I extend the model by Gohin (2006) and my paper contributes by including both bindingness and capitalization. In my model I illustrate rigorously the argument that bindingness and capitalization may lead to subsidies having negative effects on the output level, and non-bindingness and partial capitalization may lead to positive production effects. I take cattle payments under CAP 1992 as a case for binding subsidies, and those under Agenda 2000 for non-binding ones. Both CAP 1992 and Agenda 2000 are designed to reduce over-production in agriculture in the EU. Binding cattle payments under CAP 1992 can be appropriate for this objective, whereas non-binding payments under Agenda 2000 may go counter to reducing over-production.

I empirically assess production effects of cattle payments. I take advantage of the availability of Spanish farm-level data by the European Farm Accounting Data Network (FADN).[1] The data are representative and available for the period under both policy programs. Real output of beef-cattle is measured by market sales of cattle and the size of the cattle is adjusted according to the European official standard. I specify both a static and dynamic model for panel data. Fixed Effects and different GMM estimators are used. I compare the estimated effect of direct payments under CAP 1992 and under Agenda 2000. I define different farm groups for robustness checks, depending on whether a farm relies on the market for inputs (thus subsidies are likely to capitalize), and whether the payments are likely to be binding or non-binding.

Without capitalization, the "pure" effect should be zero for binding subsidies and be positive for non-binding ones. In my estimations, the estimated effect includes both a "pure" effect and, presumably, a negative effect of higher input prices due to capitalization. Negative impacts of cattle payments on beef-cattle outputs are captured under CAP 1992. Under Agenda 2000, the estimated effect of cattle payments turns to be positive. Moreover, the estimated positive effect under Agenda 2000 is statistically different from that under CAP 1992. In robustness checks, estimated effects are negative under CAP 1992 for specially defined farm groups, for whom bindingness and capitalization are more likely to happen. Estimated effects are positive under Agenda 2000 for specially defined groups, for whom the two characteristics are less likely to take

[1] I study the case of Spain where CAP 1992 takes place between 1994 and 1999 and Agenda 2000 is implemented between 2000 and 2006. Spain is the fifth largest beef producer in Europe. Beef and veal account for around 6% of the value of total agricultural products and around 20% of the value of the livestock products from 2000 to 2003. Subsidies on beef-cattle occupy a large portion of total subsidies on livestocks. This portion is about 40% from 1995 to 1999 and increases to a value between 50% to 60% from 2000 to 2003.

For the application of my paper to other countries, I can't think of any reason why the estimation results I get could not be replicated with the data from other Member States of the EU.

place.

The rest of this paper is organized as follows. Section 2 describes the CAP policies and presents a model to study the effect of cattle payments on beef-cattle outputs under each policy program. Section 3 describes the data and summarizes relevant variables. Section 4 presents the econometric models and estimation results. Section 5 gives robustness checks. Section 6 concludes.

2. Policies and Modelling Framework

2.1. Policy Background

The CAP in the 1960s provides a price support system, mainly based on intervention prices, to agricultural production. Intervention prices are set by the European Commission (EC) that is obliged to buy in surplus produce when the internal market price falls below the intervention price. This price support system leads to massive excess supply in the 1980s and, in order to reduce the over-production, the EU carries out the CAP 1992 reforms, which are intended to bring intervention prices closer to world market prices and enforce a new system of direct payments to compensate for farmers' income loss. Agenda 2000 is an extended continuation of CAP 1992. Under Agenda 2000, intervention prices are further cut down and pre-existing direct payments are revised.

Within the beef and veal sector, under Agenda 2000 intervention prices are gradually cut down from January 2000 to June 2002 by 20%. Since July 2002, the former intervention system is replaced by a new one under which both intervention prices and private storage aids are applied. Intervention prices decrease to 45% of the original level in 2000. To prevent a sharp decrease in the market price, private storage aids are granted to the farmers when the market price is likely to remain at a basic price, which is set at 66% of the intervention price in 2000.

Hans van Meijl et. al. (2002) pointed out that beef market prices have historically been on par with intervention prices. The difference between the market price and the intervention price is reversely related with the extent of the excess supply. Even after the reduction in intervention prices under Agenda 2000, there continued to be massive excess supply in the cattle-beef sector, hence, market prices should be close to intervention prices.

Under both sets of policies, there are two kinds of direct payments. One is issued on the basis of the heads of cattle raised by a holding and is referred to as headage payments.[2] The other kind doesn't depend on the heads of cattle and it includes payments to farmers in areas with environmental restrictions, compensatory allowance in less-favored areas, subsidies to rural development, etc. The latter kind will be referred to as other payments in the rest of this paper.

For headage payments, potentially eligible cattle are those kept on a holding and are referred to as cattle stocks. The amount of payment per head varies according to the type of the cattle. Headage payments include Beef Special Premium (BSP), Suckler Cow Premium (SCP) and Extensification Payment (EP). All these payments are subject to stock density limits measured as the number of livestock units per hectare (LUs per ha).[3] The stock density limit is 2 LUs per ha for BSP and SCP. EP is an additional payment on the other two payments whose stock density limit is 1.4 or 1 LU per ha for EP. The lower the density, the higher the payment per head.

Holdings whose stock density is greater than 1.4 LUs per ha are not eligible for EP. The maximum number of eligible cattle in LUs of a holding in each year is obtained by multiplying the area of land by the density limit of 2 LUs per ha. This maximum number of eligible cattle can be seen as farm-level quotas. Farmers can raise more than one type of cattle. If a holding's stock density is between 1.4 and 2 LUs per ha, its cattle stock is lower than the number of quotas. It is profitable for the holding to raise more cattle until reaching the stock density of 2 LUs per ha.

Under Agenda 2000, pre-existing headage payments are gradually increased from 2000 to 2002 and maintained at the level of 2002 until the Single Payment Scheme (SPS) came into effect in Spain in 2006.[4] Stock density limits remain the same for BSP and SCP while they become stricter for EP. A new payment, Slaughter Premium (SP), is introduced at the start of 2000, making all beef-cattle eligible as long as the cattle are killed and sold in the market. There is a regional ceiling set at the country level to limit the maximum number of eligible cattle for SP. If overall claims made by farmers exceed the regional ceiling, the eligible cattle per holding will be reduced proportionately. Since farmers can not predict if there will be over-claims when they

[2]A holding is all the production units (farms) managed by a producer. A producer could be an individual, a company or partnership.

[3] Following the European official regulation, LUs are defined as follows: 0.2 LU for calves younger than 6 months, 0.4 LU for cattle between 6 months and 1 year, 0.7 LU for cattle between 1 and 2 years, 1.0 LU for cattle older than 2 years and 1.0 LU for diary and suckler cows.

[4]The SPS is announced in June 2003 by the EC. Direct payments are decoupled from production by the SPS in the sense that direct payments only depend on historial payments of reference years and production activity is not a requirement for receiving the payments.

claim for SP, it is reasonable to assume that they will claim on all potentially eligible cattle. If there are over-claims, the payment of SP per cattle declines. Thus, one can think of SP as a payment without farm-level quotas where the payment per cattle varies depending on the incidence and the extent of over-claims.[5]

2.2. Modelling of Cattle Payments

Previous studies generally don't model the unsubsidized cattle explicitly. Instead, they assume that farmers take into account an average SCP and BSP for all their animals (Binfield et al. (2005)). The marginal animal (and hence the herd size) is determined by the equalization of price and marginal cost less the average SCP or BSP. Thus, removing or reducing the SCP will negatively affect the output.

As pointed out by Gohin (2006), from available statistics, nearly 22% of suckler cows and 20% of bulls and steers are not subsidized in 15 EU Member States. Given this fact, Gohin (2006) provides an alternative modelling framework by which the presence of unsubsidized cattle matters with respect to production effects of cattle payments. He sets up two model specifications about farmers' production decisions distinguishing whether there are unsubsidized cattle. He models all types of cattle and related cattle payments in the same way and focuses on suckler cow activity. He concludes that when SCP is limited by quotas and there are unsubsidized suckler cows, the marginal animal (and hence herd size) is determined by the standard equalization condition between price and marginal cost. SCP may have very limited production effect and simply generates rent to quota owners. On the contrary, when all suckler cows are subsidized, SCP have positive production effects.

I extend the model by Gohin (2006), which will be referred to as the initial model hereafter, by incorporating several issues as follows. First, capitalization of cattle payments is taken into account. The inclusion of capitalization alters conclusions of the initial model about production effects of cattle payments. Second, Gohin (2006) doesn't model SP initiated under Agenda 2000. Third, stock density requirements on both SCP and BSP function as quota limits as described in Section 2.1. These requirements add extra source of bindingness in cattle payments. The inclusion of stock density requirements doesn't change conclusions of the initial model. Fourth, since the payment per animal is different between SCP and BSP and farmers can produce two types of cattle at the same time, it is reasonable to argue that farmers have the incentive to raise the animal that receive more payment per head.

Suppose there is a representative farmer operating on a piece of land. As mentioned in Section 2.1, market prices are close to intervention prices. Following Gohin (2006), I denoted one price for the output by p.

[5] Detailed rules for each headage payment are available in the Appendix.

Consider the situation without direct payments and denote the output by n^0. The profit of the farmer is the following,

$$\pi = pn^0 - c(n^0)$$

where $c(\cdot)$ is the total cost function satisfying $c'(\cdot) > 0$ and $c''(\cdot) < 0$.

First order condition with respect to the output is

$$p = c'(n^0)$$

2.2.1. The Initial Model

The farmer needs the same number of quota certificates to obtain the same units of SCP and the certificates are given exogenously.[6] Denote the number of quota certificates by $\hat{q}$.

Case 1: Payments are binding, i.e. $\hat{q} \leq n^0$. Denote the payment per head by s and total payments the farmer receivesare $S = s\hat{q}$. The profit is

$$\pi = pn + s\hat{q} - c(n)$$

First order condition is

$$p = c'(n)$$

The output implied by the above condition is the same as that without direct payments, i.e. $n = n^0$. There is no production effect of binding payments. This case is relevant with cattle payments under CAP 1992.

According to the rules by the EC, quota certificates of SCP can be traded on the market. Denote the price of the quota certificate per unit by p_q. Suppose the farmer owns initially q_0 units of certificates and finally can have q units through the market and total payments are $S = sq$. q depends on the availability of quota certificates on the market.

Suppose that the payments are binding, i.e. $q \leq n^0$. The farmer needs to choose the ouput level, n, and the number of quota certificates, q, to maximize its profit,

$$\pi = pn + sq - c(n) - p_q q + p_q q_0$$

First order conditions are,

$$p = c'(n)$$

[6] To distinguish from quotas due to stock density requirements, quotas of SCP studied by Gohin(2006) are referred to as quota certificates in this paper.

$s = p_q$

The first equation implies that the output is the same as that without payments, i.e. $n = n^0$. The second one indicates that the marginal benefit of obtaining one more unit of quota certificate equals the marginal cost.

Case 2: Payments are non-binding and all units produced are subsidized. Total payments are $S = sn$. The profit is

$\pi = pn + sn - c(n)$

First order condition is

$p = c'(n) - s$

The output implied by the above condition is greater than that without direct payments, i.e. $n > n^0$. This case is relevant with cattle payments under Agenda 2000 when SP is introduced. As argued in Section 2.1, SP is non-binding. s can be thought of the average payment per animal.

2.2.2. Inclusion of Capitalization

I extend the initial model by including capitalization of the payments. Capitalization leads to an increase in input prices, such as labor prices, capital rents and feeding costs that affect the variable cost. All units produced are equally affected by the capitalization. Denote the effect of capitalization on the profit by σn, where σ is the increase in input prices. Moreover, σ is assumed to be positively related with the total payment, i.e. $\sigma = \sigma(S)$ such that $\sigma_S^{'} > 0$. Assume that capitalization does not exceed total payments, i.e. $S \geq \sigma n$.

Case 1: Payments are binding, i.e. $\hat{q} \leq n^0$, the profit is

$\pi = pn + s\hat{q} - c(n) - \sigma n$

First order condition is

$p = c'(n) + \sigma$

Comparing with the situation without payments, if bindingness and capitalization occur, direct payments negatively affect the output, i.e. $n < n^0$. This case is relevant with cattle payments under CAP 1992.

If quota certificates are traded on the market, suppose that the payments are binding, i.e. $q \leq n^0$, and total payments are $S = sq$. The profit is,

$\pi = pn + sq - c(n) - \sigma n - p_q q + p_q q_0$

First order conditions are,

$p = c'(n) + \sigma$

$s = p_q + \sigma'(S) * s * n$

The first equation implies that the output is lower than that without payments, i.e. $n < n^0$. The second one indicates that the marginal benefit of obtaining one more unit of quota certificate equals the marginal cost, which is the sum of the price of quota certificate and the resulting capitalization due to greater total payments.

Case 2: Payments are non-binding and total payments are $S = sn$. The profit is

$\pi = pn + sn - c(n) - \sigma n$

First order condition is

$p = c'(n) + \sigma - s$

If capitalization is partial, i.e. $S > \sigma n$ or $s > \sigma$, the output will be greater than that without direct payments. This case is relevant with cattle payments under Agenda 2000.

2.2.3. Incorporating Stock Density Requirements

As mentioned in Section 2.1, stock density limits define the maximum number of eligible cattle for direct payments. Denote this maximum number (or quotas) by $\bar{q}$. The farmer doesn't need to obtain more quota certificates than the limits, i.e. $q \leq \bar{q}$.

Case 1: Payments are binding, i.e. $q \leq \bar{q} < n^0$ or $q \leq n^0 < \bar{q}$.

The analysis of this case is the same as Case 1 in Section 2.2.2.

Case 2: Under Agenda 2000, SP is not limited by stock density requirements and payments are non-binding. The analysis of this case is the same as Case 2 in Section 2.2.2.

2.2.4. Two types of Cattle Payments

The farmer can apply for both SCP and BSP and the farmer needs to raise two types of cattle accordingly. I name the cattle that are eligible for SCP as type 1 cattle and those eligible for BSP as type 2 cattle.[7] Denote the output of type 1 cattle without direct payments by n_1^0, and n_2^0 for type 2 cattle.

The sum of two types of eligible cattle cannot exceed the stock density limit, $\bar{q}$. To receive BSP, the farmer doesn't need any certificate. Denote the payment per head of SCP by s^1 and by s^2 for BSP, such that $s^1 > s^2$.

[7] Suckler cows and calves are treated as a combination and referred to as type 1 cattle. The output of type 1 cattle is the sum of suckler cows and calves. So are the production costs.

The farmer needs to choose ouput levels for both types of cattle and the number of quota certificates, q, to maximize its profit.

Case 1: $\bar{q} < n_1^0$, $\bar{q} - q < n_2^0$. Both types of payments are binding. This case is relevant with cattle payments under CAP 1992.

For the moment, assume that the farmer is able to obtain the profit-maximizing number of quota certificates through the market. Total payments are $S = qs^1 + (\bar{q} - q)s^2$ or $q(s^1 - s^2) + \bar{q}s^2$. Denote the effect of capitalization on the profit through type 1 cattle by $\sigma n_1$and by σn_2 through type 2 cattle. Assume that capitalization does not exceed the total payment net the cost of quota certificates, i.e. $S - p_q q \geq \sigma n_1 + \sigma n_2$. The profit is

$$\pi = pn_1 + pn_2 + S - c_1(n_1) - c_2(n_2) - \sigma n_1 - \sigma n_2 - p_q q + p_q q_0$$

where $c_i(\cdot)$ is the total cost function pf type i cattle satisfying $c_i'(\cdot) > 0$ and $c_i''(\cdot) < 0$, with $i = 1, 2$.

First order conditions are the following,

$$p = c_1'(n_1) + \sigma$$

$$p = c_2'(n_2) + \sigma$$

$$s^1 - s^2 = p_q + (n_1 + n_2)\sigma_S'(s^1 - s^2)$$

The first two equations imply that the output of each type of cattle is lower than that without direct payments, i.e. $n_i < n_i^0$, $i = 1, 2$. The third equation implies that the marginal benefit of having one more unit of quota certificate equals the related marginal cost. The marginal cost is the sum of the price of the quota certificate and the increase in the capitalization.

In reality, there is a maximum number of animals at the national level that can receive SCP. Farmers may not be able to obtain as many certificates as they want. The availability of quota certificates can vary across farmers. Holding other conditions unchanged, the more certificates a farmer can get, the greater is total payments, the greater is σ, and the lower the output of each type of cattle. Hence, a negative relation between total payments and the output (n_1, n_2 or $n^1 + n^2$) can be conjectured.

Case 2: $\bar{q} > n_1^0$, $\bar{q} - q < n_2^0$. BSP is always binding. This case is relevant with cattle payments under CAP 1992.

Case 2.1: Suppose the farmer can not obtain more quota certificates than n_1^0, i.e. $q < n_1^0$ (SCP is binding), then the analysis of this case is the same as Case 1 of Section 2.2.4.

Case 2.2: Suppose the farmer is able to obtain more quota certificates than n_1^0, i.e. $n_1^0 < q \leq \bar{q}$ (SCP is non-binding). It is not necessary to keep more certificates than the output level, i.e. $n_1 = q$. The payment is non-binding for type 1 cattle, but binding for type 2 cattle. Total payments are $S = n_1 s^1 + (\bar{q} - n_1)s^2$ or $n_1(s^1 - s^2) + \bar{q}s^2$. The profit is

$$\pi = pn_1 + pn_2 + S - c_1(n_1) - c_2(n_2) - \sigma n_1 - \sigma n_2 - p_q n_1 + p_q q_0$$

The first order conditions are,

$$p = c_1'(n_1) + \sigma + p_q + ((n_1 + n_2)\sigma_S' - 1)(s^1 - s^2)$$

$$p = c_2\text{'}(n_2) + \sigma$$

Above two equations mean that the output of type 1 cattle is uncertain compared with n_1^0, while the output of type 2 cattle is lower than n_2^0.

By looking at the data, it is impossible to differentiate between Case 1 and Case 2. Nevertheless, Case 1 is more likely to happen if intervention prices are high enough.

Case 3: SP is introduced under Agenda 2000. Consider the case when both SCP and BSP are binding, i.e.$\bar{q} < n_1^0$ and $\bar{q} - q < n_2^0$.

Denote the amount of SP per head by s^{sp}. Total payments are $S = qs^1 + (\bar{q} - q)s^2 + (n_1 + n_2)s^{sp}$ or $q(s^1 - s^2) + \bar{q}s^2 + (n_1 + n_2)s^{sp}$. The profit is,

$$\pi = pn_1 + pn_2 + S - c_1(n_1) - c_2(n_2) - \sigma n_1 - \sigma n_2 - p_q n_1 + p_q q_0$$

If the profit-maximizing level of quota certificates of SCP are accessible, then first order conditions are the following

$$p = c_1'(n_1) + \sigma + ((n_1 + n_2)\sigma_S' - 1)s^{sp}$$

$$p = c_2'(n_2) + \sigma + ((n_1 + n_2)\sigma_S' - 1)s^{sp}$$

$$(s^1 - s^2) - (n_1 + n_2)\sigma_S'(s^1 - s^2) - p_q = 0$$

From the third equation, we have $(n_1 + n_2)\sigma_S' - 1 = \frac{p_q}{s^2 - s^1} < 0$. The output of either type of cattle can be greater than the level without direct payments, i.e.$n_1 > n_1^0$, $n_2 > n_2^0$, if $\sigma + ((n_1 + n_2)\sigma_S' - 1)s^{sp} < 0$.

If the profit-maximizing level of quota certificates of SCP are not accessible, total payments will decline. Then,

$$(s^1 - s^2) - (n_1 + n_2)\sigma_S'(s^1 - s^2) - p_q < 0 \text{ or } (n_1 + n_2)\sigma_S' - 1 > \frac{p_q}{s^2 - s^1}$$

The output, both n_1 and n_2, will decrease compared with the situation when the profit-maximizing level of quota certificates of SCP are accessible. Hence, a positive relation between total payments and the output (n_1, n_2 or $n^1 + n^2$) can be conjectured.

2.3. Empirical Identification Strategy

Headage payments are paid to eligible cattle stocks. Implementation rules of headage payments affect cattle output indirectly measured as market sales. BSP is paid to eligible bulls or steers once in the lifetime of the cattle, so that farmers have incentives to sell the cattle after receiving the payments. SCP is paid to eligible suckler cows once a year. On the one hand, old suckler cows need to be replaced with young ones. On the other hand, suckler cows are kept for the purpose of raising calves for meat production and are not allowed to provide milk to the market. Hence, SCP affects the output of calves.

Other payments are not coupled to the heads of cattle and should not affect the cattle production directly. Nevertheless, all types of direct payments can capitalize into input prices. Therefore, other payments should be taken into account in the analysis of the relation between subsidies and cattle output.

It is reasonable to argue that under both policy programs farmers tend to raise more cattle than the number of quotas when facing intervention prices. If this is the case, payments are binding under CAP 1992. This assumption can be checked with the data by comparing the actual output and the heads of cattle that actually receive direct payments, though the data are only available from 2000 on. This issue about the data will be discussed in Section 5.2.

The identification strategy of my paper is based on the assumption that the availability of SCP quota certificates for each holding is exogenous. Thus, the variation in total payments is exogenous. Since latent outputs without direct payments n_1^0 and n_2^0 are not observable, the identification strategy of this paper is to look at the change in total direct payments over time and the change in observed outputs over time.

I conduct several robustness checks. First, with available data I compare the actual output and the cattle that actually receive direct payments. Second, I expect that for farms whose inputs reply on the market supply to a greater extent, their direct payments are more likely to capitalize into input prices. The degree of capitalization depends on the competitiveness of the inputs markets. A negative effect of binding direct payments on outputs is more likely to be captured. Third, there are farms with a relatively high stock density, implying that they produce at a level above the number of quotas to a greater extent. Again, a negative effect of binding direct payments is more likely to be captured. Some characteristics of the holdings related with robustness checks are summarized in Table 1 of Section 3.1.

3. The Data

The data are from the Spanish section of the European Farm Accounting Data Network (FADN), namely the Red Contable Agraria Nacional (RECAN). This survey is an annual farm survey conducted by the Spanish

Ministry of Agriculture. The questionnaire is filled in by accountancy agencies that collect information directly from the commercial farms. A commercial farm is defined as a farm that is large enough to provide a main activity for the farmer with a level of income sufficient to support his/her family. Since 1985, the RECAN survey has been part of the European FADN. Importantly, since 1988 it has been conducted with only minor methodological changes in the definition of livestock products. Although the panel is unbalanced, most of the farms are present in the survey for several years.[8]

3.1. Sample Description and Related Issues

I want to estimate the effect of cattle payments on beef-cattle outputs under both CAP 1992 and Agenda 2000. Farm-level data from 1995 to 2003 are used. The sample studied includes holdings that raise beef-cattle and receive direct payments during this period. I consider four periods. The period under CAP 1992 includes 1995-1998. The announcement period of Agenda 2000 is from May 1999 to the end of 1999. The transitional period of Agenda 2000 refers to 2000 and 2001 when intervention prices are relatively high and headage payments are gradually increased. The full-implementation period of Agenda 2000 includes 2002 and 2003, as new policies are finally established in July 2002.[9] Year 1999 is dropped from the analysis, as it partially belongs to the period under CAP 1992 and partially to the announcement period of Agenda 2000.

Farms enter and exit the sample every year for unreported reasons. If farms enter or exit the sample independently of the implementation of new policies under Agenda 2000, cross sectional data is a representative sample and provides reliable estimates for the population of interest. Given that revised direct payments are designed to partially compensate for farmers' income loss, reduced intervention prices may cause farms with low productivity to exit the business. If this is the case, selection problems with the cross sectional data will bias the estimates for Agenda 2000.

[8] In the original dataset, the farm identifier was reused to identify another farm two years after the original farm dropped out of the sample. To avoid identification problems, in this paper it is assumed that farms that dropped out the sample would not enter the sample again.

[9] Single Payment Scheme (SPS) is another important reform of the CAP which is announced in June 2003 and started in 2006 in Spain. Since the reference amount of single payment is fixed at the level before the announcement, farmers can not react strategically to be eligible for more payments after the announcement and before the implementation.

Balanced longitudinal data are a counter-example to the original sample. Farms that exit or newly enter the business after the new policies are excluded.[10] However, if farms stay in business after the new policies are with high productivity, the balanced sample also suffers selection bias. I report estimation results with both the full and the balanced sample. If the results are similar between the two samples, estimates are less likely to be biased.

Given a common intervention price for all holdings, a holding with high productivity or low production cost will has a stock density level greater than others with low productivity. Table 1 presents the computed stock density for the full and the balanced sample. The stock density is computed by dividing cattle stocks measured in LUs by the land area of a holding.[11]

(Insert Table 1 around here.)

The mean and standard deviations in parentheses of the computed stock density are reported in columns (2) and (5) of Table 1, while column (3) and (6) give the percent of holdings with a computed stock density greater than 1 LU per ha. Holdings from the balanced sample have greater stock density than those from the full sample, especially from 2000 to 2003. This indicates selection problems with the balanced sample.

The dataset provides accurate information about cattle types for cattle stocks in all years, but not for market sales before the year 2000. More precisely, diary cows can not be exclude from the output variable before 2000 and they are not eligible for headage payments. I include sales of diary cows under both policies to maintain the consistency in the measure of outputs. At the time, in order to be eligible for headage payments under both sets of policies, the number of diary cows that farmers can raise depends on the quantity of milk quotas they hold. Since milk quota is not reformed throughout the period studied in this paper, the effect of measurement error can be alleviated by taking time difference in the output variable.

Table 2 gives a summary of diary cow stocks. Stocks are relevant since diary cows are kept for the purpose of producing milk for the market.

(Insert Table 2 around here.)

---

[10] Balanced longitudinal data which includes holdings present in the sample from 1995 to 2003 will be referred to as the balanced sample hereafter. The repeated cross sectional data which includes all holdings present in the sample in each year will be referred to as the full sample hereafter.

[11] Since the dataset doesn´t provide exact information about how much land of a holding is devoted to cattle-raising, the computed stock density potentially underestimates the actual stock density.

Column (2) for the full sample suggests that there is a time trend in the evolution of diary cow stock from 1995 to 2003. This is also true for the balanced sample in column (6). On average, holdings from the full sample keep more diary cows than those from the balanced sample in all the years, but the difference is small. Looking at column (4) for the full sample, the percentage of holdings that raise dairy cows remains relatively stable between 60% and 70% except in 1996, 1999 and 2003. Column (8) for the balanced sample shows that this percentage is quite close between the two samples and slightly greater with the full sample in most of the years.

In short, there is no clear evidence of substitutions between beef-cattle and diary cows after the announcement of Agenda 2000, either with the full or the balanced sample. Since the time trend presents with both types of sample, the issue of including diary cows in the output variable is equally important for both samples. Taking time difference and including time dummies in regressions can presumably solve this problem.

3.2. Summary of Key Variables of the Estimation Sample

Beef-cattle outputs are measured by market sales. The average sales per holding from 2000 to 2003 is summarized in columns (2) and (3) of Table 3, measured in LUs and market value separately. The value per LUs is computed by dividing the aggregate market value of the outputs of all holdings by aggregate outputs in LUs and is reported in column (4) of Table 3. Panel A of the table contains statistics from the full sample and Panel B includes those from the balanced sample defined in Section 3.1.

(Insert Table 3 around here.)

Column (2) in Panel A shows that the average sales per holding measured in LUs increases in 2001. Potential reasons for this trend can be the Agenda 2000 reforms, the recovery from the BSE crisis and an increase in beef consumption because of the reduction in intervention prices. At the same time, the average market value per holding decreases in 2001, which likely reflects the cut in intervention prices. Lower intervention prices are also reflected in the value per LU. All these features are also present in the balanced sample.

The increase in average sales in 2002 can be due to several factors, such as increased direct payments, the currency reform of Euros and an increase in consumption because of lower market prices. If the composition of types of cattle remained the same, the value per LUs in column (4) should have declined due to the cut in intervention prices in 2002. However, this doesn't seem to be the case. Note that the collected sample varies over years and, as a result, the composition of different types of cattle can vary in the sample. Moreover, the biggest cut in intervention prices takes place in July 2002 but the statistics in Table 3 are annual figures.

We can assess the importance of changes in the composition of cattle over years by looking at the balanced sample in Panel B, since the composition of cattle in the balanced sample should be more stable than that of the full sample. Column (2) in Panel B shows that in 2002 the average sales also increase as with the full sample. Interestingly, the value per LU in column (4) does decrease, which is consistent with the big cut in intervention prices in 2002.

Table 4 summarizes direct payments on beef-cattle from 1995 to 2003. Figures in panel A are obtained from the full sample and those in panel B are from the balanced sample. Headage payments are the sum of all types of headage payments. Overall payments include all types of direct payments, including headage payments and other payments defined in Section 2.1.

(Insert Table 4 around here.)

Looking at Panel A of Table 4, we can see that aggregate headage payments in column (2) keep on increasing from 1995 to 2003. Average headage payments in column (3) keep on increasing from 1996 to 1999, but decrease in 2000. Arguably, the BSE crisis negatively influences the number of cattle raised by a holding. After that, average headage payments keep on increasing from 2001 to 2003. All of these facts are also true for the overall payments in columns (4) and (5).

Figures in panel B show that aggregate and average headage payments in columns (2) and (3) keep on increasing from 1996 to 2003. This is also true for the overall payments in columns (4) and (5). Moreover, average headage payments are greater with the balanced sample than those with the full sample in all the years. This is also true for overall payments.

Table 5 gives a summary about each type of headage payment. Average payment per holding and the computed payment per cattle are reported. The payment per cattle is computed by dividing the total amount of each type of payment by the number of cattle that receive the payment. The full sample is used. The computed payment per cattle can be compared with the official values in Table 15 in the Appendix.

(Insert Table 5 around here.)

For SCP and BSP, the average payment depends on the number of eligible cattle and the payment per cattle. As shown in Table 5, the average payment keeps on increasing from 2000 to 2003. This is consistent with increases in the payment per cattle reported in Table 15 in the Appendix.

The payment per cattle of SP for cattle under seven months old is much lower than the official value (50 euros per cattle) in 2000 and 2001. This is also true for SP for cattle above eight months old whose official value is 80 euros per cattle. These facts probably show certain evidence of over-claims for SP. If claims made by farmers for SP exceed the regional ceiling, as a result, the payment per eligible cattle declines. If over-claims also happen with BSP, farmers have incentives to sell young cattle instead of keeping them for future BSP. Farmers have incentives to claim on all eligible cattle for SP. This will reinforce the severeness of over-claims for SP. The payment per cattle increases in 2002 and 2003 and becomes very close to the official value in 2003.

The computed payment per head of SCP for suckler cows is quite close to the official values shown in Table 15, whereas that of BSP for bulls is smaller than the official values, which may again imply the incidence of over-claims.[12]

Regarding the EP, the average payment depends on the number of eligible cattle and the payment per cattle. Both the average payment per holding and the payment per cattle are quite stable throughout the four years.

By looking at figures about beef-cattle sales in Table 3 and direct payments in Table 4, it is impossible to tell the relation between the payments and the cattle output, as there are factors other than direct payments that affect the output level. We need to run regressions to obtain the effect of direct payments on beef-cattle outputs while controlling for factors that affect the output and the amount of payments at the same time, such as farm size and geographic location of the farm. In addition, variables that are arguably determined simultaneously with the farm size and also influence the output are included in regressions, namely working hours devoted to the farmland, costs on machinery, costs exclusively spent on cattle raising, other intermediate costs.

One potential endogeneity problem is that if there are heterogeneous shocks affecting direct payments and productivity (thus outputs) at the same time, the identification strategy of this paper may result in biased estimates. For instance, changes in personal abilities of the farm manager may affect outputs and total payments in the same direction. In this case, the estimations give conservative estimates under CAP 1992, but overestimate production effects of direct payments under Agenda 2000. Nevertheless, the period studied in this paper is quite short and it seems reasonable to expect that unobservable factors such as personal abilities don't change much.

Control variables are summarized in Table 6. The mean and the number of observations in parentheses are presented.

(Insert Table 6 around here.)

[12]BSP for steers and SCP for heifers are not reported in Table 5, since the number of observations are very small (around 30 or fewer). The information is available upon request.

Variable “Rented Land” is the area of farmland that is rented by the farmer, while “Total Land” is the overall area of farmland utilised by a holding. According to the figures in Table 6, there are around 55% to 70% farmers rent farmland. “Paid Wage” is the total amount of wage paid to farm workers. “Unpaid Hours” is the number of working hours that are not paid, “Paid Hours” is the working hours that are paid, and “Total Hours” is the sum of these two terms. Between 10% to 20% holdings hire farm workers. “Costs on Mach” is the costs spent on renting, maintaining or renewing of machinery. “Costs on Cattle” is the costs exclusively spent on raising cattle. As is shown in Table 6, “Costs on cattle” represents an important source of costs for the holdings. “Inter Costs” are intermediate costs, including water, electricity and insurrance.

4. Econometric Models and Estimation Results

4.1. Econometric Models

4.1.1. A Static Model

As mentioned in Section 3.1, three periods are included in the regressions: the period under CAP 1992 (1995-1998), the transitional period (2000-2001) and the full-implementation period of Agenda 2000 (2002-2003).

Let t denote different years and let l denote different periods, namely $l = l_0\ if\ 1995 \leq t \leq 1998$, $l = l_1\ if\ t = 2000, 2001$ and $l = l_2\ if\ t = 2002, 2003$. The index i indicates individual holdings which raise cattle and receive direct payments. Consider the following static model for panel data:

$$y_{it} = \varphi_l S_{it} + \phi_l' Z_{it} + \theta_t + \eta_i + v_{it} \tag{1}$$

where y_{it} is the log of market sales of beef-cattle measured in LUs, S_{it} is the log of direct payments, and vector Z_{it} includes the log of the control variables that are summarized in Table 6. Time dummies, θ_t, capture the effects of changes in the economic environment such as declines in intervention prices, the BSE crisis, the Euro reforms, and imports and exports shocks. The time-invariant holding-specific term, η_i, captures time-invariant heterogeneity among holdings, such as the productivity of farm land and personal abilities of the farm manager.

The coefficient φ_l contains both the direct payment and the capitalization effect. When both bindingness and capitalization happen, φ_l is expected be negative. When payments are non-binding and capitalization is partial, φ_l is expected to be positive. The vector of coefficients $\{\varphi_{l_0}, \varphi_{l_1}, \varphi_{l_2}\}$ is the interest of this paper.

The fixed effet, η_i, is assumed to be (strictly) exogenous.

$$E(\eta_i \cdot v_{it}) = 0,\ for\, \forall t = 1, ...T. \tag{2}$$

Taking first differences in Equation 1, we have

$$\Delta y_{it} = \varphi_l \Delta S_{it} + \phi_l' \Delta Z_{it} + \Delta\theta_t + \Delta v_{it} \tag{3}$$

If direct payments, S_{it}, and other controls, Z_{it}, are (strictly) exogenous, i.e. uncorrelated with past, present and future values of v_{it}, OLS estimator can provide unbiased estimates of φ_l and ϕ_l'. However, S_{it} and Z_{it} are likely to be correlated with past shocks. For instance, if unanticipated technological shocks enable a holding to raise more of the cattle that are eligible for greater payment per head, total direct payments would increase from then on. Consequently, production methods would change, hence the input variables included in Z_{it}. Hence, I assume S_{it} and Z_{it} to be weakly exogenous (or predetermined).

$$E(S_{i,t-k} \cdot v_{it}) = 0,\ E(Z_{i,t-k} \cdot v_{it}) = 0,\ for\, t = 1, ...T,\ k \geq 0. \tag{4}$$

If there is no serial autocorrelation in v_{it},

$$Cov(v_{it}, v_{is}) = 0,\ where\, t \neq s \tag{5}$$

unbiased estimates of φ_l and ϕ_l' can be obtained using instrumental variable estimation, i.e. lagged values $S_{i,t-k}$ and $Z_{i,t-k}$ with $k \geq 1$ are valid instruments for ΔS_{it} and ΔZ_{it} respectively.

To see if φ_l varies significantly in different periods, I specify a flexible model. For simplicity, consider a model with two periods, l_0 and l_1. Equation 3 can be reparameterized as

$$\Delta y_{it} = \Delta\theta_t + \tau D_t^p + \varphi_{l_0} \Delta S_{it} + \gamma \Delta S_{it} D_t^p + \phi_{l_0}' \Delta Z_{it} + \zeta' \Delta Z_{it} D_t^p + \Delta u_{it} \tag{6}$$

where D_t^p is an indicator, i.e. $D_t^p = 1$ if t belongs to period l_1 and $D_t^p = 0$ otherwise, time dummies, $\Delta\theta_t$, contain the effect of changes in the economic environment within period l_0, $\tau + \Delta\theta_t$ captures the effect of changes in the economic environment within period l_1, parameter γ is the difference$\varphi_{l_1} - \varphi_{l_0}$ that measures how the effect of direct payments changes from period l_0 to l_1, the vector of coefficient ζ' is $\phi_{l_1}' - \phi_{l_0}'$, and

Δu_{it} is Δv_{it}.[13] Equation 6 can be expanded to include more than two periods, and can be applied to different subperiods within each policy program.

4.1.2. A Dynamic Model

As pointed out by Just (2003), evidence shows that weather is serially correlated which can generate serially correlated yields. Literatures in agricultural research oftens finds serially correlated shocks in the production model (Guan (2006)). In addition, adjusting costs in agricultural production make the overall results of a production plan cover a period of more than one year. As a result, a dynamic model is appropriate.

Following Arellano-Bond (1991), consider a dynamic model based on a first-order autoregressive process as follows:

$$y_{it} = \alpha y_{i,t-1} + \varphi_l S_{it} + \phi_l' Z_{it} + \theta_t + \eta_i + v_{it} \tag{7}$$

where the variables and the parameters are the same as those in Equation 1.[14] It is reasonable to think that an unpredictable shock is uncorrelated with past market sales, thus, y_{it} is assumed to be predetermined, i.e. y_{it} depends on past and present values of v_{it}. If there is no serial correlation as defined in Equation 5, the model implies the following moment restrictions,

$$E(y_{i,t-k} \cdot \Delta v_{it}) = 0,\ for\, t = 3,...T,\ k \geq 2. \tag{8}$$

which means that values of y lagged two periods or more are valid instruments for Equation 7 in first differences. In general, if v_{it} is MA(q) in the following sense,

$$E(v_{it} v_{i,t-s}) \neq 0,\ for\, s \leq q \tag{9}$$

moment restrictions become,

$$E(y_{i,t-q-k} \cdot \Delta v_{it}) = 0,\ for\, t = (q+3),...T,\ k \geq 2. \tag{10}$$

[13] If there is any time-invariant bias in the estimate of φ_l due to any failure of the exogeneity conditions, we can still obtain an unbiased estimate of $\varphi_{l_1} - \varphi_{l_0}$.

[14] A second order autoregression model is tried with the data of 1995-1998 and of 2000-2003, however, the estimated coefficient of $y_{i,t-2}$ is not significantly different from zero. Estimation results are available upon request. I stick to a first-order autoregression model in this paper.

I assume S_{it} and Z_{it} to be predetermined as defined in Equation 4. If v_{it} is MA(q) with $q \geq 0$, $S_{i,t-q-k}$ and $Z_{i,t-q-k}$ with $k \geq 1$ are valid instruments for ΔS_{it} and ΔZ_{it} respectively. To test for serial correlations of v_{it}, I use m_1 and m_2 statistics by Arellano-Bond (1991). Test results will be reported together with estimation results.

The estimator exploiting Equation 10 is referred to as GMM-FD in this paper. [15]

Similar to Equation 6, Model 7 admits a reparameterization,

$$\Delta y_{it} = \Delta\theta_t + \tau D_t^p + \alpha\Delta y_{i,t-1} + \varphi_{l_0}\Delta S_{it}$$

$$+\gamma\Delta S_{it}D_t^p + \phi_{l_0}'\Delta Z_{it} + \zeta'\Delta Z_{it}D_t^p + \Delta u_{it} \tag{13}$$

where τ , φ_{l_0}, γ , ϕ_{l_0}', ζ' and u_{it} are the same as those in Equation 6.

As mentioned in Section 3, farm-level data are collected by agencies directly from farmers. Although there are main guidelines for collecting data, agencies may differ, if only marginally, in their implementation of the accounting conventions set up within the FADN statistical framework. In view of possibly correlated errors of farms with the same agency, cluster-robust standard errors are captured by controlling for agencies.[16]

[15] There can be weak-instruments problems with GMM-FD, especially when the value of α increases towards unit or as the variance of η_i increases. To solve this problem, Arellano-Bover (1995) and Blundell-Bond (1998) proposed to exploit moment conditions under the assumption of no serial correlation as follows:

$$E(\Delta y_{i,t-s}\cdot v_{it}) = 0,\ for\, t = 3,...T,\ s \geq 1. \tag{11}$$

If we combine Equation 11 with an additional assumption, $E(\Delta y_{i,t-s}\cdot\eta_i) = 0$, we have:

$$E(\Delta y_{i,t-s}\cdot(\eta_i + v_{it})) = 0,\ for\, t = 3,...T,\ s \geq 1. \tag{12}$$

The GMM estimator exploiting both Equation 8 and 12 are usually referred to as System-GMM. The moments of Equation 12 exploits that the first observation comes from the stationary distribution. This may not be the case in my paper, as the institutional settings are changing and some periods should be required for farms to adapt to new settings. In practice, I use Sargan difference test by Arellano-Bond (1991) to check the validity of moment restrictions of Equation 12. Under both CAP 1992 and Agenda 2000, I find that Sargan difference test does't support the use of additional restrictions of of Equation 12.

[16] Most farms in the sample had the same agencies throughout the three periods. However, about 5% of the holdings change their agencies at some point between 1995 and 1998, and about 6% between 2000 and 2003. For these farms, agencies are chosen to be those in the first year of a given period.

4.2. Estimation Results

According to the analysis in Section 2, without capitalization, binding subsidies under CAP 1992 should have no effect on outputs, while non-binding subsidies under Agenda 2000 should have positive effects. With capitalization, binding subsidies can negatively influence outputs, whereas non-binding subsidies can have positive (zero) impacts if the capitalization is partial (full). Non-binding subsidies may even negatively affect outputs if capitalization overtakes subsidies. In this section, I provide estimation results separately for each policy program, and I compare the effect of direct payments between the two programs.

4.2.1. Effects of Direct Payments under CAP 1992

As farmers and input providers need time to react to the policies of CAP 1992, I split l_0 into two subperiods, namely $h = h_0 \; if \; 1995 \leq t \leq 1996$ and $h = h_1 \; if \; 1997 \leq t \leq 1998$. I estimate Equation 6 of the static model using First Differences (FD). Direct payments and other controls are treated as predetermined, i.e. $S_{i,t-k}$ and $Z_{i,t-k}$ with $k \geq 1$ are used as instruments for ΔS_{it} and ΔZ_{it} respectively. Estimation results for both the unconditional and conditional models are reported in Table 7. Both types of sample are used.

(Insert Table 7 around here.)

As is shown in the table, m_1 and m_2 statistics in the four columns indicate the residuals are serially correlated, which indicates that a dynamic model should be used. The serially correlated residuals don't characterize the error term of the dynamic model. Even if the error term of the dynamic model is MA(0), the wrongly specified static model would present serially correlated residuals as shown in the table.

I estimate Equation 7 of the dynamic model using FD-GMM. Since the data are available only for four years, lagged values of y_{it} as instruments for $y_{i,t-1}$ in first differences are available if the error term is MA(q) with $0 \leq q \leq 1$. I start with estimating the equation by assuming the error to be MA(1). Since the order of serial correlation of the residuals cannot be identified with m_1 statistic alone, I use Sargan difference test by Arellano-Bond (1991) to indirectly determine the validity of different lags as instruments. In the end, estimation results using $y_{i,t-3}$ as a instrument for $\Delta y_{i,t-1}$ and $S_{i,t-k}$ with $k \geq 2$ for ΔS_{it} are reported in Table 8.

(Insert Table 8 around here.)

Columns (1) and (2) contains estimates with the full sample, while columns (3) and (4) for the balanced sample. In the four columns, the estimated coefficient, $\hat{\alpha}$, is negative and within the unit circle. It is only significant with the full sample. In previous empirical studies about the agricultural outputs, the sign of the estimated coefficient of the lagged dependent variable shows a mixed picure. Some find a estimate between 0 and 1, e.g. Mythili (2008) and Kanwar (2008), while others find a negative one, e.g. Yu et al. (2012) and Kanwar (2008).

Estimates of S in columns (1) and (2) show that direct payments have negative impacts on outputs during 1997-1998, but they are not significant. The negative impacts mean that with a 1% increase in direct payments, outputs would decrease by around 0.5%. Looking at columns (3) and (4), the estimated effect of direct payments is larger in magnitude than that of the full sample. Moreover, the effect is statistically significant, implying that with a 1% increase in direct payments, outputs would decrease by around 1.2% during 1997-1998.

4.2.2. Effects of Direct Payments under Agenda 2000

As defined in Section 3.1, the period under Agenda 2000 is divided into two periods, i.e. the transitional period l_0 and the full-implementation period l_1. I estimate Equation 6 of the static model using FD. Results are presented in Table 9.

(Insert Table 9 around here.)

Similar to Table 7, m_1 and m_2 statistics in the four columns indicate the residuals are serially correlated, implying that a dynamic model is needed.

I estimate Equation 7 of the dynamic model using FD-GMM. Similar to the situation in Section 4.2.1, the data are available only for four years, i.e. 2000-2003. I apply the same method to determine the validity of different lags as instruments. Finally, estimation results using $y_{i,t-3}$ as a instrument for $\Delta y_{i,t-1}$ and $S_{i,t-k}$ with $k \geq 2$ for ΔS_{it} are reported in Table 10.

(Insert Table 10 around here.)

Similar to the results in Table 8, the estimated coefficient, $\hat{\alpha}$, is negative and within the unit circle in the four columns.

Looking at columns (1) and (2) for the full sample, estimates of S show that direct payments have positive and maginally significant effects on outputs. If direct payments increase by 1%, the output would increase by about 0.16% during 2002-2003. Estimates of S in columns (3) and (4) of the balanced sample are also maginally significant and greater than those of the full sample. That is, if direct payments increase by 1%, the output would increase by about 0.55% during 2002-2003.

4.2.3. Comparison of Two Policy Programs

In this subsection I want to check if the effect of direct payments differ between CAP 1992 and Agenda 2000. I estimate the flexible specifications defined in Equation 13 using FD-GMM. Then, I test if the estimated coefficient of direct payments differs between the two policy programs. P-values of the tests are reported in Table 11.

(Insert Table 11 around here.)

Results in columns (1) and (2) are for the full sample. Looking at either the unconditional or the conditional model, p-values show that the estimated effect of direct payments differs statistically between 1997-1998 and 2002-2003. Results in columns (3) and (4) are for the balanced sample. They also show that estimates of direct payments under Agenda 2000 are different from those under CAP 1992 at a significance level of 10%.

4.2.4. Decomposition of Policy Effects

How much of the change in the output, $\bar{y}_{il_2} - \bar{y}_{il_0}$, is due to the new policies can be written as $\hat{\varphi}_{l_2} * \bar{S}_{il_2} - \hat{\varphi}_{l_0} * \bar{S}_{il_0}$. I decompose it into two parts as follows,

$$\hat{\varphi}_{l_0} * (\bar{S}_{il_2} - \bar{S}_{il_0}) + (\hat{\varphi}_{l_2} - \hat{\varphi}_{l_0}) * \bar{S}_{il_2} \tag{14}$$

where the first term measures the counterfactual effect of an increase in payments under Agenda 2000 while φ_l is maintained at the level under CAP 1992, and the second term measures the counterfactual effect of a change in φ_l due to new rules of Agenda 2000 while the payments are fixed at the level of Agenda 2000.

Decomposition results using the full sample are reported in Table 12. Estimates of direct payments are obtained using the dynamic model reported in Tables 8 and 10. Estimates of direct payments during 1997-1998 are used to represent $\hat{\varphi}_{l_0}$, while those during 2002-2003 to represent $\hat{\varphi}_{l_2}$.

(Insert Table 12 around here.)

Columns with the label "Payments" give the results for the first term of Equation 14 and show that the percentage increase in direct payments would lead to a decrease in output because $\hat{\varphi}_{l_0}$ is negative. Columns with the label "Policy" give the results for the second term, i.e. the impact on outputs due to changes in the elasticity between direct payments and outputs. Since $\hat{\varphi}_{l_2}$ during 2002-2003 are positive in all specifications, the second term in Equation 14 is positive and captures the output effect of the policy change from CAP 1992 to Agenda 2000 even if subsidies are unchanged.

Results in column (3) of the conditional model with the full sample show that the increase in direct payments would cause outputs to drop under CAP 1992 by 8.2% during 2002-2003. Estimates in column (7) of the balanced sample show that outputs would decline by 19.1% during 2002-2003. Slightly smaller estimates are obtained with the unconditional model as shown in columns (1) and (5).

5. Robustness Checks

Under CAP 1992, if a holding has unsubsidized outputs, it means the subsidies are binding. If a holding relies on market for input supplies, it is more likely to suffer capitalization of subsidies. In this section, I explore whether a negative association between direct payments and outputs can be capture for these holdings. On the contrary, I also want to check whether a positive association can be obtained for holdings that are less likely to suffer from capitalization under Agenda 2000.

I estimate Equation 13 using data from 1995 to 1998 for different groups. Estimation method is the same as that for Table 8. Results are reported in Table 13.[17]

(Insert Table 13 around here.)

Estimated effects of direct payments under Agenda 2000 for different groups are reported in Table 14. Estimation method is the same as that for Table 10.

(Insert Table 14 around here.)

[17]Only results with conditional models and the full sample are reported in Tables 13 and 14. Results with unconditional models and the balanced sample are available upon request. Results with the unconditional model and with the balanced sample are quite similar to those in Tables 13 and 14.

Holdings Producing Above Quotas. As mentioned in Section 2.4, the assumption that the output of beef-cattle with direct payments is greater than the number of quotas can be checked. Assume that the heads of cattle that actually receive payments are equal to the number of quotas as long as the output is greater than the number of quotas. The question turns to be the comparison between the output and the heads of cattle that actually receive direct payments. I name those holdings with cattle sales greater than the cattle that actually receive headage payments within a year as "holdings producing above quotas". The data about the cattle that actully receive payments are available from 2000 on. To circumvent this problem, I assume that holdings that produce above quotas from 2000 to 2003 produced above quotas during 1995 to 1998.[18]

The estimated effect of direct payments under CAP 1992 for holdings producing above quotas is reported in column (1) of Table 13. The estimate of S is negative but not significant during 1997-1998.

Holdings with High Stock Density. As mentioned in Section 2.3, holdings with relatively high stock density are likely to produce above the number of quotas and, hence, subsidies are likely to be binding. If capitalization happens, subsidies have negative impacts on outputs.

Estimated effects of direct payments under CAP 1992 for holdings with stock density above 2 and 3 are reported in columns (2) and (3) of Table 13 respectively. In both columns, the estimate of S is negative and insignificant. The estimate in column (2) is quite small in magnitude.

Estimated effects of direct payments under Agenda 2000 for holdings with stock density below 3 and 4 are reported in columns (1) and (2) of Table 14 respectively.[19] Results in column (1) show that the estimated effect is positive and marginally significant during 2002-2003. The estimate is also positive in column (2), but not significant.

Holdings Not Receiving EP. According to the eligibility rules for EP, not receiving EP is a sign of high stock density and, thus, subsidies are likely to be binding. The information about whether a holding receives EP is available from 2000. To circumvent this problem, I assume that holdings that don't receive EP from 2000 to 2003 didn't receive EP from 1995 to 1998.[20] These holdings are referred to as "holdings not receiving EP"

[18]Even if the ouput is above the number of quotas, it can be that during a calendar year market sales are below the heads of cattle that actually receive payments. For instance, cattle may die or consumed by the producer afer receiving payments and before being sold on the market. It can also be that cattle are to be sold in the near future but the payments are already issued in the current year. To be consevative, only holdings with cattle sales greater than the cattle that actually receive payments are used in the estimation.

For the validity of the assumption, more than 95% of holdings, that produce above quotas in 2000, produce above quotas in 2001. The same holds from 2001 to 2002, and from 2002 to 2003.

[19]The limits of stock density (2 and 3 for CAP 1992, and 3 and 4 for Agenda 2000) are chosen to obtain sufficiently large samples.

[20]For the validity of the assumption, more than 95% of holdings, that don't receive EP in 2000, don't receive EP in 2001. The same holds from 2001 to 2002, and from 2002 to 2003.

hereafter.

Estimated effects of direct payments under CAP 1992 for holdings not receiving EP are reported in column (4) of Table 13. The estimated effect of direct payments is negative but not significant.

Holdings Relying on Markets for Inputs. Holdings whose inputs reply on the market supply to a greater extent are more likely to suffer from capitalization. Given this, different groups of holdings are defined here, including holdings that hire farm workers and pay salaries, that rent machinery, and that rent land (although land rent doesn't affect the variable cost of a holding, it may affect the costs of home-produced feedings).

Estimated effects of direct payments under CAP 1992 for holdings hiring workers, renting machinery and renting land are reported in columns (5), (6) and (7) of Table 13 separately. Estimate of S in column (5) is positive but in quite small magnitude, which is probably due to the small sample size. Estimates of S in columns (6) and (7) are negative and significant in column (7).

The estimated effects under Agenda 2000 for holdings not hiring workers, not renting machinery and not renting land are reported in columns (3), (4) and (5) of Table 14 respectively. The estimates of S in the three columns are positive and marginally significant.

Holdings with High/Low Cattle-Ratio. Besides beef-cattle, holdings can produce other crops, such as cereal, protein crops, vegetables, fruits and other kinds of livestock. I compute for each holding the ratio of the market value of beef-cattle outputs relative to that of other products. This ratio will be referred to as cattle-ratio. A holding with a relatively high cattle-ratio is more likely to rely on the market for cattle feedings and suffer from capitalization.

Estimated effects of direct payments under CAP 1992 for holdings with a cattle-ratio above 20 and 200 are reported in columns (8) and (9) of Table 13 respectively. The estimated effect of direct payments shown is negative in both columns.

The estimated effect of direct payments under Agenda 2000 for holdings with a cattle-ratio below 500 are reported in column (6) of Table 14.[21] The estimated effect is positive but not significant.

Among the nine groups in Table 13, the smallest estimate of S happens with holdings with stock density above 3 as well as holdings with a cattle-ratio above 20. The estimate indicates that if direct payments increase by 1%, outputs would decrease by around 0.8% during 1997-1998. The estimate is also relatively small for holdings renting machinery and with a cattle-ratio above 200.

[21] The limits of cattle-ratio (20 and 200 for CAP 1992 and 500 for Agenda 2000) are chosen to obtain sufficiently large samples.

Among the six groups in Table 14, the greatest estimate of direct payments happens with holdings not renting machinery, indicating that if direct payments increase by 1%, outputs would increase by about 0.5% during 2002-2003. Other groups are quite similar in the estimated effect of direct payments.

6. Conclusions

I study production effects of subsidies taking into account both bindingness and capitalization. Without capitalization, binding subsidies have no production effect, whereas non-binding subsidies have positive effects. With capitalization, bindingness leads to subsidies having a negative production effect, while non-bindingness may cause a positive, zero or negative effect depending on the extent of capitalization. I study the case of cattle payments implemented in Spain using a panel dataset that is representative and includes more than 1000 observations in each year. Under CAP 1992, eligibility rules on cattle payments and intervention prices for beef and veal make it likely that payments are binding. Under Agenda 2000, cattle payments become non-binding.

Estimation results suggest that cattle payments under CAP 1992 have negative impacts on outputs during 1995-1996 as well as during 1997-1998. The estimated effects of cattle payments are highly significant during 1997-1998. Under Agenda 2000, the estimated effect is positive and marginally significant. The imprecision in the estimations may be caused by the limited length of the panel data used.

The decompostion analysis of policy effects gives estimates of the counterfactual effect of an increase in payments in Agenda 2000 period under the CAP 1992 regime. In my preferred model specification, results show that on average the increase in direct payments during 2002-2003 would cause outputs to decline by 8.2% under the CAP 1992 regime. After restricting the sample to holdings that present in the dataset from 1995 to 2003, results show that outputs would decline by 19.1%.

Results from robustness checks show that, when cattle payments are likely to be binding and holdings count on markets for input supplies (i.e. capitalization is likely to happen), the estimated effect of the payments is negative. When cattle payments are non-binding and holdings don't reply on markets for inputs (i.e. capitalization is less likely to happen), the estimated effect is positive.

CAP 1992 and Agenda 2000 reforms are both designed to reduce the over-production by cutting down intervention prices and compensating farmers' income loss with direct payments. The intention of introducing direct payments is not to increase the production. Within the beef and veal sector, estimates found in this paper indicate that cattle payments are negatively associated with beef-cattle outputs under CAP 1992. However, this association becomes positive under Agenda 2000. The effect goes counter to the objective of Agenda 2000 to reduce the over-production.

Table 1: Comparison of the Full and the Balanced Sample

BASIC STATISTICS

	Full Sample			Balance Sample		
	Obs. (1)	Density (2)	% of density>1 (3)	Obs. (4)	Density (5)	% of density>1 (6)
1995	955	30.3 (200.8)	59%	182	38.2 (256.8)	65%
1996	1337	28.6 (218.7)	55%	376	36.1 (212.5)	64%
1997	1791	22.1 (187.8)	50%	558	25.1 (173.7)	62%
1998	1476	28.1 (205.9)	60%	558	26.5 (176.2)	64%
1999	1389	28.4 (209.7)	60%	446	32.7 (197.1)	61%
2000	1976	17.1 (150.3)	56%	459	36.2 (206.3)	66%
2001	1944	17.9 (156.9)	57%	465	35.4 (207.3)	65%
2002	1558	23.6 (168.8)	62%	489	42.8 (233.5)	67%
2003	1650	25.1 (178.6)	64%	529	42.5 (236.7)	66%

Note: "Density" is the computed stock density. "% of density>1" is the percent of holdings with a computed stock density greater than 1 LU per ha.
"% Machine" is the percent of holdings that use machinery.
"% Workers" is the percent of holdings that hire workers.

Table 2: Average Stocks per Holding of Diary Cows

BASIC STATISTICS

	Full Sample				Balance Sample			
	Obs. (1)	Mean (2)	Std. (3)	% (4)	Obs. (5)	Mean (6)	Std. (7)	% (8)
1995	656	22.7	17.5	69%	122	19.2	15.4	67%
1996	1017	25.3	17.6	76%	283	22.9	13.5	75%
1997	1315	25.3	22.3	73%	375	25.3	15.6	67%
1998	926	26.1	18.8	63%	367	24.9	14.8	66%
1999	779	31.7	23.0	56%	247	25.1	16.6	55%
2000	1358	32.4	22.0	69%	260	28.2	16.7	57%
2001	1277	34.1	23.7	66%	263	31.1	19.5	57%
2002	958	37.9	26.7	61%	290	33.2	22.2	59%
2003	932	41.0	39.8	56%	324	34.0	22.8	61%

Note: "Obs." is the number of holdings that raise diary cows. "Mean" is the average heads of diary cows per holding. "Std" is the standard errors. "%" is the percentage of holdings raising diary cows out of the holdings that raise beef-cattle and receive direct payments.

Table 3: Cattle Output Measured by LUs and Market Value

Basic Statistics

Panel A: Full Sample

	Obs. (1)	LUs Mean (2)	Value Mean (3)	Value per LU (4)
2000	2004	14.5	15,709	1,085
2001	1965	18.3	15,161	829
2002	1563	20.6	18,977	921
2003	1652	22.8	24,795	1,089

Panel B: Balance Sample

	Obs. (1)	LUs Mean (2)	Value Mean (3)	Value per LU (4)
2000	460	14.8	19,915	1,346
2001	465	16.3	19,251	1,184
2002	489	23.3	22,232	954
2003	529	18.1	23,473	1,298

Note: "LUs" is market sales measured in LUs, "Value" is market value of cattle sales in euros. "Value per LU" is the computed value per LU.

Table 4: Summary of Direct Payments on Cattle

BASIC STATISTICS

Panel A: Full Sample

		Headage Payments		Overall Payments	
	Obs.	Aggregate	Mean	Aggregate	Mean
	(1)	(2)	(3)	(4)	(5)
1995	715	2,314,527	3,237.1	2,457,956	3,437,7
1996	1,071	3,431,377	3,203.9	3,676,850	3,433.1
1997	1,147	4,117,397	3,589.7	4,374,314	3,813.7
1998	1,152	4,513,421	3,917.9	4,845,335	4,206.0
1999	1,218	5,943,718	4,879.9	6,164,785	5,061.4
2000	1,849	7,882,381	4,263.1	8,350,565	4,516.3
2001	1,563	8,373,147	5,357.1	9,650,525	6,174.4
2002	1,526	9,097,234	5,961.5	10,753,569	7,046.9
2003	1,654	13,782,782	6,333.0	15,317,553	7,260.9

Panel B: Balanced Sample

		Headage Payments		Overall Payments	
	Obs.	Aggregate	Mean	Aggregate	Mean
	(1)	(2)	(3)	(4)	(5)
1995	118	454,996	3,855,9	488,515	4,140,0
1996	267	862,116	3,228.9	975,778	3,654.6
1997	380	1,716,992	4,518.4	1,837,095	4,834.5
1998	392	1,746,674	4,455.8	1,914,528	4,884.0
1999	349	1,809,251	5,184.1	1,935,648	5,546.3
2000	386	2,305,308	5,972.3	2,525,814	6,543.6
2001	325	2,549,365	7,844.2	2,823,607	8,688.0
2002	470	2,921,849	6,216.7	3,473,587	7,390.6
2003	521	3,536,055	6,787.1	4,010,033	7,696.8

Note: "Headage Payments" is the sum of all types of headage payments in euros under CAP 1992 and under Agenda 2000. "Overall payments" is the sum of "Headage Payments" and other payments. All payments are measured in euros.

Table 5: Summary of Different Headage Payments

BASIC STATISTICS

	SP < 7m			SP > 8m			BSP for bulls		
	Obs	Avg. Pay.	Pay.per Head	Obs	Avg. Pay.	Pay.per Head	Obs	Avg. Pay.	Pay.per Head
2000	858	743	22.7	595	948	35.4	225	2,911	125.6
2001	657	476	17.3	732	1,010	33.4	325	3,823	141.6
2002	327	981	39.6	1,152	1,593	63.7	334	4,167	164.3
2003	458	800	53.0	1,360	2,190	78.8	462	3,911	195.4

	SCP for sucklers cows			EP		
	Obs	Avg. Pay.	Pay.per Head	Obs	Avg. Pay.	Pay.per Head
2000	643	5,159	181.0	247	3,487	99.6
2001	681	6,125	166.1	310	3,160	89.7
2002	587	6,349	189.5	290	3,567	96.7
2003	710	7,311	185.0	374	3,967	100.5

Note: “SP<7m” (“SP>8m”) refers to Slaugher Premium for cattle younger than 7 months (older than 8 months). “Avg. Pay.” is the average payment per holding in Euros. “Pay. per Head” is the computed payment per head in Euros.

Table 6: Summary of Control Variables

BASIC STATISTICS

	1995	1996	1997	1998	1999	2000	2001	2002	2003
Rented Land	16.7	36.6	35.5	45.9	54.9	49.1	51.4	36.1	41.9
	(488)	(738)	(920)	(876)	(884)	(1113)	(1174)	(1065)	(1154)
Total Land	22.2	34.4	36.2	57.1	69.4	56.6	60.2	54.6	61.2
	(956)	(1338)	(1792)	(1480)	(1393)	(2001)	(1965)	(1563)	(1652)
Paid Wage	2533.3	4339.3	4974.9	4113.7	5884.8	5674.0	6760.7	6222.4	7895.9
	(96)	(145)	(190)	(170)	(266)	(271)	(256)	(245)	(304)
Unpaid Hours	3306.4	3242.1	3046.8	2989.0	3250.0	3307.6	3133.2	3136.2	3231.8
	(957)	(1336)	(1788)	(1429)	(1388)	(2000)	(1961)	(1559)	(1647)
Paid Hours	652.3	907.9	1082.1	945.3	979.2	1118.6	1282.6	1211.1	1287.2
	(96)	(145)	(190)	(170)	(266)	(271)	(256)	(245)	(304)
Total Hours	3368.4	3335.7	3154.6	3097.1	3425.3	3452.3	3294.0	3318.0	3458.9
	(958)	(1338)	(1792)	(1431)	(1393)	(2004)	(1965)	(1563)	(1652)
Costs on Mach	1344.3	1590.6	1971.0	2059.4	2137.9	2281.6	3293.7	3242.1	3636.9
	(880)	(1283)	(1660)	(1307)	(1220)	(1806)	(1837)	(1446)	(1517)
Costs on Cattle	16451.1	21320.0	22019.0	22599.4	28338.3	27549.6	30096.1	34959.8	39351.1
	(958)	(1338)	(1792)	(1472)	(1393)	(2003)	(1965)	(1563)	(1651)
Inter Costs	875.3	1135.1	998.9	1053.5	1692.5	1671.4	1482.3	2050.0	2313.4
	(907)	(1269)	(1599)	(1264)	(1323)	(1883)	(1770)	(1515)	(1602)

Note: The mean and the number of observations in parentheses are reported in the table. "Rented Land" is the area of farmland that is rented by the farmer. "Total Land" is the overall area of farmland utilised by a holding. Both variables are in 100 hectares. "Paid Wage" is the total amount of wage paid to farm workers. "Unpaid Hours" is the number of hours worked that are not paid, while "Paid Hours" is the hours paid for. "Total Hours" is the sum of these two terms. "Costs on Mach" is the costs spent on the renting, maintenance or renewing of the machinery. "Costs on Cattle" is the costs exclusively spent on raising cattle. "Inter Costs" is some intermediate costs including water, electricity and insurrance. All variables except the land are in Euros.

Table 7: Effects of Direct Payments under CAP 1992 (Static Model, 1995-1998)

ESTIMATED COEFFICIENTS AND STD ERROR

	Full Sample		Balanced Sample	
	Uncondi.	Condi.	Uncondi.	Condi.
	(1)	(2)	(3)	(4)
S	0.685**	0.507***	0.414	0.386
	(0.274)	(0.194)	(0.293)	(0.262)
$S * D_{1998}$	0.228	0.151	-0.112	-0.079
	(0.151)	(0.119)	(0.127)	(0.083)
D_{1996}	0.221***	0.259***	0.182***	0.359***
	(0.039)	(0.038)	(0.067)	(0.088)
D_{1997}	-0.148**	-0.100	-0.133	-0.135
	(0.075)	(0.066)	(0.129)	(0.139)
D_{1998}	0.547***	0.600***	0.730***	1.004***
	(0.095)	(0.082)	(0.145)	(0.154)
No. of Obs.	2756	2704	1038	1015
m_1	0.000	0.000	0.000	0.000
m_2	0.012	0.038	0.003	0.089
Sargan test	0.010	0.003	0.302	0.341

Note: S is overall direct payments. S and additional controls are summarized in Table 4 and Table 6. Estimates of additional controls are omitted. D_{1996} is a time dummy that equals 1 during 1996, and 0 otherwise. D_{1997} and D_{1998} are also time dummies. Robust standard errors are given in brackets. For m_1, m_2 and Sargan tests, p-values are reported. Instruments used for ΔS_{it} are $S_{i,t-k}\, for\, k \geq 1$. Specifically, instruments used for $\Delta S_{i,1998}$ are $S_{i,1997}$, $S_{i,1996}$ and $S_{i,1995}$.

*, **, *** means significant at the 90th, 95th, 99th percentile.

Table 8: Effects of Direct Payments under CAP 1992 (Dynamic Model, 1995-1998)

ESTIMATED COEFFICIENTS AND STD ERROR

	Full Sample		Balanced Sample	
	Uncondi.	Condi.	Uncondi.	Condi.
	(1)	(2)	(3)	(4)
$y_{i,t-1}$	-0.492**	-0.491**	-0.786	-0.705
	(0.232)	(0.254)	(0.609)	(0.781)
S	-0.449	-0.580	-1.167**	-1.179*
	(0.513)	(0.503)	(0.592)	(0.731)
D_{1996}	0.137	0.201	0.452***	0.477
	(0.131)	(0.132)	(0.179)	(0.305)
D_{1997}	0.648***	0.702***	0.519***	0.480***
	(0.126)	(0.145)	(0.122)	(0.126)
D_{1998}	0.712***	0.747***	0.884***	0.863***
	(0.090)	(0.110)	(0.116)	(0.158)
No. of Obs.	1584	1552	594	583
m_1	0.002	0.003	0.056	0.063
Sargan test	0.329	0.339	0.345	0.264

Note: $y_{i,t-1}$ is the lagged dependent variable. Refer to Table 7 for definitions of variables. Robust standard errors are shown in brackets. For m_1 and Sargan tests, p-values are reported. Instruments used for ΔS_{it} are $S_{i,t-k}\, for\, k \geq 2$, and for $\Delta y_{i,t-1}$ are $y_{i,t-k}\, for\, k \geq 3$. Specifically, instruments used for $\Delta S_{i,1998}$ is $S_{i,1996}$ and $S_{i,1995}$, and for $\Delta y_{i,1997}$ is $y_{i,1995}$. *, **, *** means significant at the 90th, 95th, 99th percentile.

Table 9: Effects of Direct Payments under Agenda 2000 (Static Model, 2000-2003)

ESTIMATED COEFFICIENTS AND STD ERROR

	Full Sample FD (S is pred.)		Balanced Sample FD (S is pred.)	
	Uncondi.	Condi.	Uncondi.	Condi.
	(1)	(2)	(3)	(4)
S	0.278***	0.224***	0.245***	0.232***
	(0.018)	(0.016)	(0.027)	(0.024)
$S*D_{2003}$	0.194***	0.121***	0.237***	0.135***
	(0.023)	(0.020)	(0.030)	(0.028)
D_{2001}	-0.376***	-0.331***	-0.518***	-0.398***
	(0.046)	(0.043)	(0.060)	(0.062)
D_{2002}	0.005	-0.023	0.079***	0.087***
	(0.018)	(0.018)	(0.028)	(0.030)
D_{2003}	0.172***	0.173***	0.127***	0.042
	(0.030)	(0.027)	(0.040)	(0.044)
No. of Obs.	3297	3297	1118	1118
m_1	0.001	0.002	0.000	0.001
m_2	0.000	0.000	0.001	0.000
Sargan test	0.000	0.000	0.000	0.000

Note: Refer to Table 7 for definitions of variables. D_{2001} is a time dummy that equals 1 during 2001, and 0 otherwise. D_{2002} and D_{2003} are also time dummies. Robust standard errors are shown in brackets. For m_1, m_2 and Sargan tests, p-values are reported. Instruments used for ΔS_{it} are $S_{i,t-k}\, for\, k \geq 1$. Specifically, instruments used for $\Delta S_{i,2003}$ are $S_{i,2002}$, $S_{i,2001}$ and $S_{i,2000}$.
*, **, *** means significant at the 90th, 95th, 99th percentile.

Table 10: Effects of Direct Payments under Agenda 2000 (Dynamic Model, 2000-2003)

ESTIMATED COEFFICIENTS AND STD ERROR

	Full Sample FD-GMM		Balanced Sample FD-GMM	
	Uncondi.	Condi.	Uncondi.	Condi.
	(1)	(2)	(3)	(4)
$y_{i,t-1}$	-0.163***	-0.171***	-0.118	-0.130*
	(0.050)	(0.049)	(0.082)	(0.080)
S	0.158*	0.147*	0.546*	0.584*
	(0.098)	(0.087)	(0.314)	(0.328)
D_{2001}	-0.084	-0.073	-0.623	-0.680
	(0.178)	(0.173)	(0.530)	(0.529)
D_{2002}	0.027	0.043*	0.040	0.075
	(0.028)	(0.027)	(0.056)	(0.052)
D_{2003}	0.297***	0.277***	0.143	0.149*
	(0.051)	(0.050)	(0.101)	(0.088)
No. of Obs.	1483	1483	503	503
m_1	0.002	0.001	0.011	0.015
Sargan test	0.151	0.171	0.120	0.141

Note: Refer to Tables 7 and 9 for definitions of variables. Robust standard errors are in brackets. For m_1 and Sargan tests, p-values are reported. Instruments used forΔS_{it} are $S_{i,t-k}\, for\, k \geq 2$, and for $\Delta y_{i,t-1}$ are $y_{i,t-k}$ $for\, k \geq 3$. Specifically, instruments used for $\Delta S_{i,2003}$ is $S_{i,2001}$ and $S_{i,2000}$, and for $\Delta y_{i,2002}$ is $y_{i,2000}$. *, **, *** means significant at the 90th, 95th, and 99th percentile.

Table 11: Comparison of Two Policy Programs

ESTIMATED COEFFICIENTS AND STD ERROR

	Compare 2002-2003 with 1997-1998			
	Full Sample		Balanced Sample	
	Uncondi.	Condi.	Uncondi.	Condi.
	(1)	(2)	(3)	(4)
P-values	0.014	0.013	0.073	0.064
No. of Obs.	3067	3035	1097	1086
m_1	0.002	0.003	0.016	0.022
Sargan test	0.328	0.228	0.325	0.271

Note: I test if the estimated effect of direct payments during the period under CAP 1992 and that under Agenda 2000 are different. P-values of the tests are reported.

Table 12: Decomposition of Policy Effects

ESTIMATED COEFFICIENTS AND STD ERROR

Dynamic Model							
Full Sample				Balance Sample			
Uncondi.		Condi.		Uncondi.		Condi.	
Payments	Policy	Payments	Policy	Payments	Policy	Payments	Policy
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
-6.3%	1.354	-8.2%	1.621	-18.9%	4.052	-19.1%	4.171

Note: "Payments" give the results for the first term of Equation 14, i.e. the percentage increase in direct payments leads to a percentage decrease in outputs. "Policy" give the results for the second term, i.e. the impact on outputs due to changes in the elasticity between direct payments and outputs.

Table 13: Robustness with Different Farm Groups, CAP 1992

ESTIMATED COEFFICIENTS AND STD ERROR

	Dynamic Model (Full Sample)								
	>quotas	density>2	density>3	no EP	workers	machine	land rent	cattle-ratio>20	cattle-ratio>200
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$y_{i,t-1}$	-0.302	-0.407**	-0.431	-0.554**	-0.355	-0.620**	-0.442	-0.409*	-0.492
	(0.434)	(0.212)	(0.378)	(0.262)	(0.452)	(0.255)	(0.494)	(0.238)	(0.339)
S	-0.485	-0.039	-0.811	-0.227	0.046	-0.771**	-0.457	-0.809	-0.727
	(0.746)	(0.608)	(0.729)	(0.453)	(0.332)	(0.399)	(0.454)	(0.588)	(0.579)
No. of Obs.	968	1008	613	995	137	1427	810	1102	1058
m_1	0.016	0.031	0.054	0.003	0.049	0.033	0.089	0.008	0.005
Sargan test	0.970	0.395	0.086	0.697	0.368	0.278	0.440	0.160	0.296

Note: Refer to Tables 7 and 8 for definitions of variables. Robust standard errors are in brackets. For m_1 and Sargan tests, p-values are reported. Instruments for ΔS_{it} and $\Delta y_{i,t-1}$ are the same as those described in Table 8. *, **, *** means significant at the 90th, 95th, 99th percentile respectively.

Table 14: Robustness with Different Farm Groups, Agenda 2000

Estimated coefficients and Std Error

	Dynamic Model (Full Sample)					
	density<3	density<4	no employee	no machine	no land rent	cattle-ratio<500
	(1)	(2)	(3)	(4)	(5)	(6)
$y_{i,t-1}$	-0.161***	-0.194***	-0.147***	-0.295**	-0.239***	-0.163**
	(0.059)	(0.059)	(0.059)	(0.146)	(0.058)	(0.065)
S	0.161*	0.101	0.186*	0.502**	0.192*	0.167
	(0.092)	(0.147)	(0.109)	(0.229)	(0.114)	(0.137)
No. of Obs.	1006	1191	1309	161	538	658
m_1	0.009	0.051	0.010	0.075	0.006	0.006
Sargan test	0.296	0.072	0.518	0.046	0.100	0.730

Note: Refer to Tables 7 and 9 for definitions of variables. Robust standard errors are shown in brackets. For m_1 and Sargan tests, p-values are reported. Instruments for ΔS_{it} and $\Delta y_{i,t-1}$ are the same as those described in Table 10. *, **, *** means significant at the 90th, 95th, 99th percentile.

Appendix

Regulations of each type of headage payments under Agenda 2000 are summarized in Table 15.

Table 15: Regulations of Headage Payments under Agenda 2000

	Eligible cattle	Frequency	Amount	Limits
Beef Special Premium	Bulls > 9 mths	Once in lifetime	210	A regional ceiling
	Steers : 9 and 21 mths	Twice in lifetime	150	Stock density <2
Suckler Cow Premium	Suckler cows and heifers	Annual	200	Regional & individual ceilings
				Stock density <2
Extensification Payment	Additional on BSP and SCP	Once in lifetime	40	Stock density (1.4, 1.8]
			80	Stock density (0, 1.4]
Slaughter Premium	Bulls, steers, cows > 8 mths	Once in lifetime	80	A regional ceiling
	Calves < 7 mths, < 160 kg	Once in lifetime	50	No limit on stock density

Note: The amount of headage payments were increasing gradually from 2000 to 2002 and remained at the level of 2002 until the SFP came into effect. The amount in the table is the level of 2002 in Euros. The regional ceiling for bulls and steers of Spain was 713,999 heads and for suckler cows 1,441,539 heads.

BSP is a payment for bull and steer producers. Eligible animals are bulls from the age of 9 months and steers at the age of 9 months and 21 months respectively.[22] Claims on the BSP should be submitted before the end

[22]Under CAP 1992, BSP per steer or bull was 109 euros, with an individual (holding) limit of 90 heads within each age bracket. Individual limit/ceiling was the maximum number of bulls or steers a holding could keep. At the time, the payments were defined in European currency units (ECU) which ceased to exist on 1 January 2002 in Spain and was replaced by the Euro at an exchange rate of 1:1.

Under Agenda 2000, BSP per bull was 160 euros in 2000, 185 euros in 2001 and 210 euros in 2002; BSP per steer was 122 euros in 2000, 136 euros in 2001 and 150 euros in 2002. The premium remained at the same level from 2002 to 2005. The individual limit of BSP was the same as that of CAP 1992.

of the year. There is a regional ceiling on the BSP, which is the maximum number of bulls and steers eligible for the BSP within a member state of the EU of a calendar year. The regional ceiling is fixed on the basis of the 1996 figures. If the total number of animals claimed by the farmers within a region exceeds the regional ceiling, the number of eligible animals per holding will be reduced proportionately.

The receipt of BSP is subject to a stock density limit of 2 LUs per forage hectare, with the exception of holdings having less than 15 LUs.[23] Cattle taken into account when computing the stock density include diary cows, ewes on which Sheep Annual Premium are claimed, male cattle on which BSP are claimed, and suckler cows and heifers on which Suckler Cow Premium are claimed.[24] Cattle and sheep on which no claims were made didn't count as LUs.

SCP is an annual payment.[25] Eligible animals are cows of meat producing breed kept for rearing calves for meat, but not for producing diary product for consumers. Such cows include suckler cows and breeding heifers. The eligibility for SCP is restricted by a minimum number of diary cows which are never eligible for SCP. Claims on SCP should be submitted before the end of the year. Old suckler cows can be replaced with young cattle if the old ones are to be sold or killed.

SCP is limited by both regional and individual ceilings. The regional ceiling is set at the highest level of premium payments in the years 1995, 1996 and 1997 plus 3%. The individual ceiling is equal to SCP quotas held by a holding by the end of 1999. A holding should make sure that it has enough quotas before applying. If not, the holding should buy or lease in quotas from the market, or apply for new quotas from the National Reserve. Newcomers can apply to the National Reserve for SCP quotas. The quotas obtained from the National Reserve can only be used for the purpose of applying for SCP and can't be transfered within three years after the application. The receipt of SCP is also subject to a stock density limit of 2 LUs per forage hectare. The calculation is the same as that for BSP.

EP is an additional payment per animal based on the cattle which are paid BSP or SCP. Under CAP 1992, EP is paid as long as the stock density is low enough and the calculation of stock density is the same as that for BSP and SCP.[26] However, under Agenda 2000 the calculation of stock density limit become stricter than

[23] The stock density limit requirement under CAP 1992 is also 2 LUs/ha.

[24] The Sheep Annual Premium is implemented under the CAP in November 1992 by the EU. An eligible animal is a female sheep that is either 12 months old or has given birth to a lamb. Farmers need to hold as many quotas as the number of sheep they want to claim on. A breeding ewe on which Sheep Annual Premium is claimed is counted as 0.15 LU. The premium is not revised under Agenda 2000.

[25] Under CAP 1992, SCP per animal was 145 euros. Under Agenda 2000, SCP per animal was 163 euros in 2000, 182 euros in 2001 and 200 euros in 2002. SCP remained at the same level from 2002 to 2005.

[26] Under CAP 1992, EP per animal was 36 euros if the stock density was between 1 and 1.4 LUs/ha, and 52 euros if the stock density was less than 1 LU/ha.

that for BSP and SCP. First, all cattle aged six months or above and all sheep on a holding are taken into account. Second, the forage area must include at least 50% of permanent grassland. Similarly, the lower is the stock density, the higher is the payment.[27]

SP is paid once in a lifetime of a cattle.[28] Eligible animals are calves aged from one month to seven months, and bulls, steers, cows and heifers aged over eight months. Eligible animals for a calendar year should be slaughered within the year and claims should be submitted before March of the following year. SP is not subject to stock density limit. The number of eligible animals is restrained by a regional ceiling which is determined by the regional number of cattle killed or exported in 1995. If the total number of animals claimed exceeds the regional ceiling, the number of eligible animals per holding will be reduced proportionately.

For both BSP and SCP, if the stock density limit is exceeded, eligible cattle will be reduced to meet the limit. For both BSP and SP, 60% of the payment is paid after October of the year. Any reduction in the number of eligible animals arising from exceeding the regional ceiling (over-claims) is calculated and the balance is paid between April and June of the following year. Cattle which are not paid either due to excess claims or exceeding the stock density limit will not be eligible in the following years.

The payment of SCP is also complished in two steps. 60% of the payments is made after October and the balance is made between April and June in the following year. The payment of EP is made between April and June in the following year together with the remaining balance of BSP, SCP and/or SP.

In areas designated as "less-favoured", agricultural production or activity is more difficult because of natural handicaps, e.g. difficult climatic conditions, steep slopes in mountain areas, or low soil productivity in other less favoured areas. Due to the handicap to farming there is a significant risk of agricultural land abandonment. To mitigate this risk, the Less Favoured Areas (LFA) payment scheme is an important tool. LFA payments are granted annually per hectare of utilised agricultural area. The level of the payment can vary between a minimum of 25 €/hectare and a maximum of 200 €/hectare. 57 % of the overall Utilized Agricultural Area in the EU is classified as Less Favoured Area. In 2005 approximately 1.4 million farms, representing about 13% of the total number of farms in the EU25, received support under all LFA schemes.

[27] In 2000 and 2001, EP per animal was 33 euros if the stock density was between 1.6 and 2.0 LUs/ha and 66 euros if the stock density was less than 1.6 LUs/ha. In 2002, EP per animal was 40 euros when the stock density was between 1.4 and 1.8 LUs/ha and 80 euros when the stock density was less than 1.4 LUs/ha.

[28] For cattle from the age of eight months, SP was 27 euros per animal in 2000, 53 euros in 2001 and 80 euros in 2002.

For cattle less than seven months old, SP was 17 euros per animal in 2000, 33 euros in 2001 and 50 euros in 2002. The premium remained the same level from 2002 to 2005.

[1] Abadie, A., 2000. Causual Inference in Econometrics and Program Evaluation. Harvard University, Cambridge.

[2] Arellano, M., S. Bond, 1991. Some tests of specification for panel data: Monte carlo evidence and an application to employment equations. Review of Economic Studies, 58, 277–297.

[3] Arellano, M., O. Bover, 1995. Another look at the instrumental variable estimation of error- component models. Journal of Econometrics, 68, 29–51.

[4] Binfield, J., Meyers, W. and Westhoff, P. Challenges of Incorporating EU Enlargement and CAP Reform in the GOLD Model Framework, Paper presented at the 89th EAAE Seminar, (Parma, Italy, 3–5 February, 2005).

[5] Blundell, R., Bond, S., 1998. Initial Conditions and Moment Restrictions in Dynamic Panel Data Models. Journal of Econometrics, 87, 115-143.

[6] Chinsinga, B., O'Brien, A., 2008. How Agricultural Subsidies are Working in Malawi. Africa Research Institute, ISBN No. 9781906329013.

[7] Ciaian, P, Kancs, d'A, 2012. The Capitalization of Area Payments into Farmland Rents: Micro Evidence from the New EU Member States. Canadian Agricultural Economics Society, 60, 517-540.

[8] Dunne, W., O'Connell, J.J., Shanahan, U., Drennan, M., Keane, M.G., 2009. Evaluation of Supply Control Options for Beef. Project report at Rural Economy Research Centre, Ireland.

[9] European Parliament, March 2012. Draft Report on the Future of the Common Agricultural Policy after 2013.

[10] European Commission, Directorate-General for Agriculture, 2000. Impact analyses of Agenda 2000 decisions for CAP reform.

[11] European Commission, Macsharry Reforms, Agenda 2000. Available from <http://ec.europa.eu/index_en.htm>

[12] Floyd, 1965. J. E. The Effects of Farm Price Supports on the Returns to Land and Labor in Agriculture. Journal of Political Economy, Vol. 73, No. 2, pp. 148-158

[13] Gohin, A., 2006. Assessing CAP Reform: Sensitivity of Modelling Decoupled Policies. Journal of Agricultural Economics, 57, 415-440.

[14] Guan, Z., Lansink, A. O., 2006. The Source of Productivity Growth in Dutch Agriculture: A Perspective from Finance. American Journal of Agricultural Economics, 88(3), 644-656.

[15] Gunjal K., Williams S., Romain, R., 2005. Agricultural Credit Subsidies and Farmland Values in Canada. Canadian Journal of Agricultural Economics, 44, 39-52.

[16] Hennessy, D.A., 1998. The Production Effects of Agricultural Income Support Policies under Uncertainty. American Journal of Agricultural Economics, 80, 46-57.

[17] Just, R. E., Pope, R., 2003. Agricultural Risk Analysis: Adequacy of Models, Data and issues. American Journal of Agricultural Economics, 85, 1249-1256.

[18] Kanwar, S., & Sadoulet, E., 2008. Dynamic Output Response Revisited: The Indian Cash Crops. The Developing Economies, 46(3), 217-241.

[19] Kirwan, B., 2009. The Incidence of U.S. Agricultural Subsidies on Farmland Rental Rates. Journal of Political Economy, 117, 138-164.

[20] Latruffe, L., Le Mou¨el, C., 2009. Capitalization of Government Support in Agricultural Land Prices: What Do We Know? Journal of Economic Surveys, 23, 659-691.

[21] Leeuwen, M., Tabeau, A., 2002. Impact of CAP Animal Premiums on Cattle and Ewe Stock in the Netherlands. Paper for the 10th EAAE Congress in Zaragoza, 28-31 August 2002.

[22] Mendez, J. A., Mora, R., San Juan, C., 2003. A Cointegration Analysis of the Long-Run Supply Response of Spanish Agriculture to the Common Agricultural Policy.

[23] Mythili, G., 2008. Acreage and yield response for major crops in the pre-and post-reform periods in India: A dynamic panel data approach, Mumbai: Indira Gandhi Institute of Development Research (Report prepared for IGIDR - ERS/ USDA Project: Agricultural Markets and Policy).

[24] Patton, M., Kostov, P., McErlean, S., Moss, J., 2008. Assessing the Influence of Direct Payments on the Rental Value of Agricultural Land. Food Policy, 33, 397-405.

[25] Roberts, M., Kirwan, B., Hopkins, J., 2003. The Incidence of Government Program Payments on Agricultural Land Rents: the Challenges of Identification. American Journal of Agricultural Economics, 85, 762-769.

[26] Rolph, E.R., 1952. A Theory of Excise Subsidies. The American Economic Review, Vol. 42, No. 4, 515-527.

[27] van Meijl, H., van Tongeren, F., 2002. Agenda 2000 CAP Reform, World Prices and GATT-WTO Export Constraints. European Review of Agricultural Economics, 29, 449-470.

[28] Yu, B., Liu, F., & You, L., 2012. Dynamic Agricultural Supply Response Under Economic Transformation: A Case Study of Henan, China. American Journal of Agricultural Economics, 94(2), 370-376.

Chapter 2: Reassessing The Differential Impact of Grandmothers and Grandfathers: The Old Age Program In Nepal

1 Introduction

In their recent review of the anthropological literature, Sear and Mace (2008) report that many studies find correlational evidence that maternal grandmothers tend to improve child survival rates (in around 70% of the studies they review) while paternal grandmothers show somewhat more variation in their effects on child survival. What about grandfathers? They find that the statistical association between grandfather presence and child survival is much weaker: in 10 of 12 cases, the presence of a maternal grandfather had no significant effect on child survival rates while paternal grandfathers had no effect in 6 of 12 studies.

The evidence in economic studies of a gender differential in the impact of grandparents on children's health is, to our knowledge, limited to Duflo (2000, 2003). These two studies explore the effects on children's health of the expansion of the Old Age Pension program in South Africa. They find that pensions received by women had a positive impact on the health and nutritional status of children living in the same household. When the beneficiary of the pension is a man, however, no health effects are found.

A methodological concern for the causal interpretation of these results, which is shared with those from the anthropological studies, is that conditional on a household's having three generations, the presence of an elderly grandparent may be a sign of a relatively healthy household. Duflo (2003) takes advantage of the fact that the height for age of young children depends on accumulated investments over the life of the child. Hence, if households with eligibles have worse characteristics than non–eligible households, older children would be smaller in eligible households. The identifying strategy then is to compare the difference in height between children in eligible and those in non–eligible households among children exposed to the program for a fraction of their lives to the same difference among children exposed all their lives.

In this paper, we study the effects on child mortality of the introduction in 1995 of a non-contributory universal pension scheme in Nepal known as the Old age Allowance Program, OAP. Under the OAP all Nepalese with age 75 and above were eligible to a universal flat rate pension of 100 Rupees per month, around 2 dollars and 12% of the country's income per capita. We use cross-sectional data from the 1996 and 2001 Nepal Demographic and Health Surveys (NDHS).

We first follow a standard diff-in-diffs approach to estimate the effects on infant mortality of eligibility to an exogenous increase in the income of an old female and an old male living in the same household. Our benchmark identification strategy consists of comparing the average changes in survival rates before and after the implementation of the OAP of children living in three-generation households with at most one male and one female eligibles to the OAP with four alternative population controls. Using this approach, we find positive and significant effects on survival rates for the presence of a female eligible to the OAP while negative and sometimes significant effects for the presence of a male eligible.

These results are qualitatively similar across alternative definitions of the control group. We also obtain similar results when we restrict the sample for both boys and girls. Finally, the results are robust to changing the method employed to exploit retrospective information in our data, to whether the female (male) eligbile is the only eligible in the household, and to the family status of the beneficiary.

We then conduct pre-treatment common trend tests to justify the validity of the Parallel Paths assumption in the benchmark diff-in-diffs approach and find that we cannot reject it for the case of the female eligibles but we strongly reject it for the case of male eligibles. This is consistent with a situation where endogenous composition of households together with economic progress create a downward sloping trend in the unobservable household quality on households with a male beneficiary. As this negative trend would be absent in households with only a female beneficiary of the OAP, the standard diff-in-diffs estimates would be appropriate to estimate the effect of the female beneficiary but would be inadequate to estimate the effects of the presence of a male beneficiary.

Following Mora and Reggio (2012), we propose a more flexible model and then conduct a test of pre-treatment common accelerations that would provide justification for a Parallel Growths assumption (i.e., assuming that without treatment, the change in growth for the treated would have equaled that of the controls). We cannot reject the presence of pre-treatment common accelerations for the male eligible effect and we strongly reject it for the female eligibles effect. Hence, we implement a flexible identification strategy based on the Parallel Growths assumption for the male eligible effect and on the Parallel Paths assumption for the female eligible effect. The positive effects of the female eligible effect remain similar to those obtained using the benchmark diff-in-diffs approach. In contrast, the estimates of the male eligible effect become positive and strongly significant. Thus, with a more flexible approach that standard diff-in-diffs estimation we do not find significant gender differences.

We are agnostic as to the channels through which the effects reported take place. For example, if only those eligible that worry more about the family end up collecting the benefits, then our results could not be compared with studies where only beneficiaries are studied. Finally, we argue that our results can be interpreted as suggestive that under economic growth and gender differences in the presence of a beneficiary in the household, cross-sectional analysis may bias downwards the estimates of the effect of grandfathers.

The rest of the paper is structured as follows. We first describe the institutional setting in Section 2 and then present the data and the estimation strategy in Section 3. In Section 4 we report and discuss the results of the paper. Section 7 concludes.

2 Policy background

In the last decades, Nepal has steadily ranked as one of the least developed countries in the world. In 1995, the year the OAP program was introduced, GDP per capita was 200 US dollars in real terms, ranking Nepal as the 211th country in the world. Living conditions for children were also among the worst in the world. The infant mortality rate in Nepal at the time was 7.6%, higher than the average among Asian countries (5.4%). Malnutrition incidence among children under 5 years old was 64.5% using height for age as criterion and 44.1% using weight for age, when the average for other Asian developing countries was 42.9% and 28.8%, respectively.

The OAP scheme is initially announced on December 1994 as part of a five-year economic plan. All Nepalese citizens with age 75 and above become eligible to a universal flat rate pension of 100 rupees per month, i.e. around 2 dollars or 12% of the country's real GDP per capita. There were five (out of 75) pilot districts in which the program officially started in January 1995, although the actual payments were delayed until July. During the following Nepalese fiscal year (from 16th July 1995 to 15th July 1996), the OAP is extended to the entire country.[1]

In the fiscal year 1999-2000, the government updates the OAP from 100 to 150 rupees per month (or, equivalently, from 7.3% to 11% of real GDP per capita) presumably to accommodate the pensioners' accumulated loss in purchasing power due to the large increases in nominal GDP. There were two additional rate updates since 1999. In 2005 the OAP increases from 150 to 200 rupees and in 2008 from 200 to 500 rupees, or 34% of real GDP per capita. In addition, the age threshold is reduced from 75 to 70 years old in 2008.

There are no direct measures of the actual coverage of the program. Looking at the early stages in the implementation of the program, Rajan (2003) reports that some legitimate beneficiaries may have initially found difficulties to prove both their citizenship and date of birth. Although the number of OAP recipients is relatively stable since its inception until 2001 (between $170,000$ and $175,000$), then it abruptly increases by 10%. Based on census information, Rajan (2003) estimates the coverage of OAP to be ranging in that year from 83% to 86% in 2001. Hence, if the observed increase in the number of recipients in 2001 only reflects coverage improvements, then average coverage during the first years of the implementation of the OAP may have ranged from 75% to 78%, possibly with lower coverage in poor isolated areas, where ignorance about the program was presumably larger.

[1]The Nepalese government introduces for the fiscal year 1996-1997 two additional social programs that could affect the economic conditions of the elderly. One of these two programs, the Helpless Widows Allowance, is only targeted at old widows who get neither any care from family members nor a widow pension. As we study only the income effects on children physically living with the elderly, in our data there are no individuals who can benefit from this allowance. The second social program introduced at the time, a Disabled Pension of 100 Rupees to adult disabled citizens, affected a very small proportion of the adult population. See, for example, Rajan (2003) for a more detailed description of these new policies.

3 Data and estimation strategy

3.1 The Nepal Demographic and Health Surveys

The data come from the 1996 and 2001 Nepal Demographic and Health Surveys (NDHS).[2] Each survey is divided into two questionnaires. The household questionnaire provides demographic characteristics for every member of the household—such as current age, sex, education, and relation with the Household Head—and basic information on the characteristics of the household—such as its regional location and whether it is located in a rural or urban area. The individual or woman questionnaire is targeted at women of age between 15 and 49. In addition to their birth history, demographic information—like current age, education, major occupational category, and ethnic status—for the mother (and her husband if present) are included.

An important feature of the NDHS is that, for all interviewed women between 15 and 49, it contains birth information—such as the birth date, sex, birth order, and whether the child has a twin—on all their children, regardless of whether the children are alive or dead at the time of the interview. For those children who are dead, the dataset also contains their death date. Therefore, it is possible to reconstruct monthly survival histories for all the children born from the interviewed women. It is also possible to reconstruct some retrospective information of the children when they were infants: the number of siblings they had and the age of their mothers. If the father and the grandparents live in the household at the time of the interview, it is also possible to obtain their age at the time the child was an infant.

The data have several shortcomings. First, the NDHS data does not provide information on whether old people in the survey collected the benefits. We can identify eligible individuals in the household since eligibility is only based on age, but we cannot be sure that those eligible did collect the OAP pension. Hence, our results pertain only to the effect of eligibility status. Moreover, since kin relations for each member of the household can only be reconstructed via his/her relation with the household head, we also cannot be sure that those eligible are the grand-parents of the infant.

Second, apart form birth and death dates, the data set does not contain retrospective information. This is a potential problem for those variables whose value at the time of the interview may differ from the value at the time the child was under one year old. An important example refers to the presence of grandparents in the household at the time of interview, since this presence does not imply presence at the birth of the child.[3] Hence, when we attempt to capture the effect of the presence at birth of a grandparent on infant survival by controlling for the presence at the time of interview, we potentially

[2]Both surveys are part of the worldwide Demographic and Health Survey (DHS) project. Additional information on the 2001 NDHS may be obtained from the Family Health Division, Department of Health Services, Ministry of Health, Nepal. Additional information about the DHS project may be obtained from ORC Macro (web site: http://www.measuredhs.com).

[3]Other examples include variables that may change with time, such as the presence of the father, his occupational status, and the parent's educational highest achievements.

incur in a measurement error that is likely larger the larger is the time span between the birth of the child and the time of the interview. One simple way to limit this measurement error is by restricting the estimation sample to births close enough to the interview date so that we expect that presence at the time of interview very likely implies presence at the time of birth. In the results section we present alternative restrictions of the estimation sample to discuss the sensitivity of our results to alternative assumptions on retrospective information.

Third, living (at the time of interview) infants whose mothers either do not live in the household (perhaps because they died before the interview) or are not eligible for the woman's questionnaire, cannot be included in the analysis. Although we do have their survival history, we do not have information for the corresponding population of dead children, i.e. dead infants whose mothers either do not live in the household at the time of interview or are not eligible for the woman's questionnaire. Hence, we do not include living children whose mothers do not live in the household because including them could potentially create sample selection bias.

Finally, there is no information on induced abortions. Before 2002, abortion was prohibited in Nepal and physicians could not recommend or perform it. Women seeking abortion did so clandestinely, frequently put their lives at risk, and suffered sometimes serious health or legal consequences (Thapa, 2004, Thapa and Padhye, 2001). Although there is no accurate direct information about the prevalence of abortion in Nepal at the time the OAP was implemented, the information available suggests that abortion was not a generalized method of birth control. Cross-country comparisons do show that pregnancy loss at the time was not high in Nepal (Casterline, 1989). Moreover, according to the 1996 and 2001 NDHS surveys, only around 18% of women in reproductive age reported to have had a pregnancy that terminated in a miscarriage, abortion, or still birth. Focusing on abortion-related hospital admissions, several studies show that in the last two decades of the 20th century only between 10 to 20 percent of these admissions were induced abortions (Thapa and Padhye, 2001, and the references therein). Thus, although it is impossible to know with precision the incidence of induced abortion, the available figures suggest that it has not been a generalized method of fertility control.[4]

3.2 The estimation sample

We combine two Surveys—the 1996 and the 2001 surveys—to create our data set. For the period before the government started implementation of the OAP, i.e. the pre-treatment period, we use the 1996 survey. In our benchmark estimation sample, we include all children born between July 1991 and June 1994 for the pre-treatment period. We do not include kids born between July 1994 and June 1995

[4] The law changed in 2004, effectively liberalizing abortion on several general grounds. Importantly, the new bill recognized the right to terminate a pregnancy of up to 12 weeks voluntarily. Presumably, this may have made the use of abortion as a fertility control mechanism more general. Our results, however, cannot be driven by any changes in fertility control triggered by the change in the law, as we do not use data after the law changed,

because in the pre-treatment period we want children not affected by the policy before they are one year old. We also restrict the sample to kids born at most five years before the time of the interview to limit potential measurement error regarding retrospective information. Again for this reason, we do not use observations from the 2001 survey in the pre-treatment period.

For the period after the government started implementation of the OAP—i.e. the post-treatment period—we use the 2001 survey. We include all children born between July 1995 and June 1998. We do not include observations from children born after June 1998 because, as explained in Section 2, the government updated the amount of the OAP starting in July 1999. We do not include children born between July 1995 and June 1996 from the 1996 survey because we do not know whether they survived their first year of life.

In sum, in our benchmark estimation sample we use the survival histories of all kids from the 1996 survey born between July 1991 and June 1994 for the pre-treatment period and all kids from the 2001 survey born between July 1995 and June 1998 for the post-treatment period. One nice feature of this sample is that it covers births along a span of three years in the two periods. However, the minimum gap between the birth date and the time of the interview is 1.5 years for the pre-treatment period but 2.5 years for the post-treatment period. We will come back to this issue when we review the robustness of our results to the use of alternative samples.

3.3 A difference-in-differences strategy

Consider the case of how the presence in the household of an OAP eligible individual may affect an infant's survival status one year after birth.[5] Variable survival status after one year, S, is equal to 1 if the infant still lives one year after birth and is equal to 0 otherwise. Let S^0 denote survival status in the hypothetical case that the government does not introduce the OAP and let S^1 denote survival status in case the government introduces the OAP. Additionally, let $D = 1$ if the infant lives in a household with an eligible individual and $D = 0$ otherwise. We refer to infants for whom $D = 1$ as the treated and infants for whom $D = 0$ as the controls. Potential and observed survival statuses are related by $S = S^1 D + S^0 (1 - D)$.

We follow a standard latent-variable specification for both S^1 and S^0:

$$S^v \equiv \begin{cases} 1 & \text{if } S^{*v} \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

where S^{*1} is survival score in case the government implements the OAP and S^{*0} is survival score in case the government does not implement the OAP. Both survival scores S^{*1} and S^{*0} are unobservable

[5] In the empirical application, we restrict the sample to households with at most one male eligible and one female eligible and allow for gender differences in the effects on the child survival status. For notational simplicity, we present in this section the model assuming at most only one eligible individual.

latent-variables which, together with D, drive survival status S. Note that, given equation (1), S also follows a latent variable specification:

$$S \equiv \begin{cases} 1 & \text{if } S^* \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

where $S^* = S^{*1}D + S^{*0}(1-D)$. We define the average effect on survival score for the treated as the expected change in survival score among those treated after implementation of the OAP:

$$\phi \equiv E\left[S^{*1} - S^{*0} | D = 1\right]. \tag{3}$$

The estimation of expectation $E\left[S^{*1}|D=1\right]$ is not difficult because it equals $E\left[S^*|D=1\right]$.[6] What makes identification of ϕ difficult is the identification of the average survival score for the treated in the absence of policy, $E\left[S^{*0}|D=1\right]$.

Assume we have information on survival status S before and after the start of the OAP both for infants living with and without an eligible individual. Define ΔS^{*v} as the change in survival score S^{*v} when any given infant goes from being born before the implementation of the OAP to being born after the implementation of the OAP. The parallel path assumption in this context states that, conditional on a vector of individual characteristics x, the average change in survival score in the absence of treatment is the same for treated and controls:

$$E\left[\Delta S^{*0}|D=1,x\right] = E\left[\Delta S^{*0}|D=0,x\right]. \tag{4}$$

As a result of technological progress, survival scores improve with time. In Nepal, these improvements may have been overshadowed by the negative effects of the civil unrest that started in 1996 and ended in 2006. Moreover, development failure might have been the root of the civil conflict (Sharma, 2006). Equation (4) specifies that, in case the government never implements the OAP, the average changes in survival scores are similar for the population of infants who live with an OAP eligible and the population of infants who do not live with an OAP eligible.

Average survival score levels may still differ across the treated and the controls. This would likely be the case when there is endogenous formation of households. For example, if households with OAP eligibles are statistically associated with worse economic conditions, then average survival scores will tend to be lower for the treated than for the controls. Thus, the parallel-path assumption allows for group-specific, time-invariant, unobservable heterogeneity which may arise from endogenous formation of households.

Assuming equation (4) immediately leads to a difference-in-differences moment condition for ϕ based

[6] One can, for example, assume $S^* \sim N\left(\beta_0, \sigma^2\right)$ and estimate $E\left[S^*|D=1\right]$ by ML.

on changes in survival score S^*:

$$\phi = E\left[\Delta S^* | D=1, x\right] - E\left[\Delta S^* | D=0, x\right]. \tag{5}$$

We base our benchmark identification strategy on equation (5). Assuming linearity in $E\left[S^*|D,x\right]$, condition (5) leads to a linear standard diff-in-diffs specification for survival score:

$$S^* = \beta_0 + \beta x + \gamma_D D + \gamma_P Post + \phi D \times Post + \varepsilon \tag{6}$$

where $E\left[\varepsilon|D,x\right] = 0$ and *Post* is a dummy variable for birth after the implementation of the OAP. Under the assumption of normality for error term ε we have a standard probit specification for the conditional expectation of observed survival status S:

$$Pr\left(S=1|D,x\right) = \Phi\left(\beta_0 + \beta x + \gamma_D D + \gamma_P Post + \phi D \times Post\right). \tag{7}$$

Consistent estimation by ML estimation of parameter ϕ in equation (7) gives a consistent estimate of the causal effect of the policy implementation on the survival score for the treated. To asses how this increase in the survival score affects survival probabilities, we focus our results on the average marginal effect of a score increase of size ϕ :

$$\alpha \equiv E\left[\Phi\left(\beta_0 + \beta x + \gamma_D D + \gamma_P Post + \phi\right) - \Phi\left(\beta_0 + \beta x + \gamma_D D + \gamma_P Post\right)\right] \tag{8}$$

3.4 Alternative control groups

To ensure a simple and tractable definition of treatment, we do not include in our analysis those infants who live with more than one eligible woman or with more than one eligible man. We only consider two types of treatments. An infant receives the first type of treatment if there is an eligible woman–i.e. a woman who is older than 75 at the time the infant is born–in the same household. An infant receives the second type of treatment if there is an eligible man.[7] Accordingly, there are three types of treated infants: those who live with an eligible woman, those who live with an eligible man, and those who live both with an eligible woman and an eligible man. As shown in Table 1, almost 100 infants both before and after treatment are treated.

Although controls and treated may differ in levels in a typical diff-in-diffs setup, the usual parallel-path assumption still imposes group homogeneity in pre-treatment dynamics, so it is reasonable to look for controls that are as similar as possible to the treated. We consider four alternative control groups. The first control group—that we refer to as *control 1 infants*—includes all infants who do not live with

[7] With complete retrospective information, we could separate those infants who have lived since birth with OAP eligibles from those who have only lived with OAP eligibles during a fraction of their first year. However, we can only look at children who live with OAP eligibles at the time of the interview.

Table 1: *Number of controls and treated (all observations)*

Control	Pre-treatment	Post-treatment	All
Treated infants	99	98	197
Control 1 infants	3424	4068	7492
Control 2 infants	598	680	1278
Control 3 infants	1965	2341	4306
Control 4 infants	487	546	1033

Note: Control 1 infants are infants who do not live with either an eligible woman or an eligible man. Control 2 infants are infants who live with people who were between 60 and 74 at the infant's birth date. Controls 3 infants are infants who do not live with people older than 60 in households where the household head is not older than 40 years of age. Control 4 infants are infants who live with non-eligible old people who were between 60 and 69 at the infant's birth date.

either an eligible woman or an eligible man. Control 1 infants are a very large group because it includes households with old people (defining old people as those older than 60 at the infant's birth date) and households without old people.

An alternative control group—that we refer to as *control 2 infants*—would restrict the comparison to infants who live with old people who are still non-eligible, i.e. infants who live with people who were between 60 and 74 at the infant's birth date. The number of control 2 infants is around 600 both before and after treatment. Given that the age eligibility limit, 74, was fixed by the government without considering how old people may help infants, control 2 infants are an interesting control group. One potential problem for control 2 infants, however, is that, with the introduction of the OAP program, old people who will soon become eligible may choose to increase their contributions to the household even before they become eligible because their permanent income raises with the announcement of the program. Hence, infants living in households where non-eligibles will soon become eligibles may not be convincing controls as they might actually receive benefits similar to those received by the treated.

One way to avoid the problem that arises with permanent income increases among old non-eligibles is to consider as controls those infants who do not live with people older than 60. To make this control group as homogeneous as possible, we additionally impose that the household head is not older than 40 years of age. Hence, most control 3 infants are newborns living in two-generation households.[8] Clearly, among the control 3 infants the introduction of the OAP program does not lead to an increase in the household available resources. However, in the presence of endogeneity in the member composition of the households, it can be argued that control 3 infants can be less appealing as a control group than control 2 infants. For example, suppose that young couples attempt to live in their own houses as soon as they reach a minimum income. As economic conditions improve nationwide, economic conditions in households with grandparents will worsen relative to households without grandparents and this trend differential will make the standard parallel-paths assumption inappropriate. Finally, grandparents who do not live in the household may still live close enough to have an influence in the welfare of the infant, so that control 3 infants we may have infants that could be considered to be under a weak version of treatment.

[8]Note that the sum of the number of control 2 and 3 infants is actually smaller than the number of control 1 infants because in the latter group there are also infants who live only with their non-eligible grand-parents.

A less radical way to avoid the permanent income increase problem is to consider as control group only infants who live with non-eligible old people younger than 69. We refer to them as *control 4 infants*. Control 4 infants are a subset of control 2 infants, the sample size being, unsurprisingly, the smallest among the four control definitions (487 before treatment and 546 after treatment). Admittedly, there could still be an increase in the household resources driven by the elderly expectations to receive the benefits in the future. We believe, however, that this effect should be smaller than the effect for control 2 infants for two complementary reasons. First, for the elderly living in households of control 4 infants, the minimum time interval before any benefits are obtained is five years. Second, life expectancy at 60 for the 1995 to 2000 period was around 16 years for women and 15 years for men.[9] Hence, a large proportion of non-eligible elderly between 60 and 69 do not survive to become eligible.

4 Diff-in-diffs results

4.1 The basic diff-in-diffs estimates

In Table 2 we report the estimated average marginal effects as defined by equation (8) under the four alternative control groups. We report the p-values for the significance tests of the estimates in parenthesis. We allow for different effects for male and female eligible and present two specifications. The unconditional specification allows for month of birth and region fixed effects.[10] In addition to the fixed effects included in the unconditional specification, the conditional specification includes two dummy variables for the education of the mother (for primary, for secondary, and for higher education), the mother's age at the child's birth and its square, a dummy for whether the child is female, the number of kids younger than five—at birth of infant—in the household, and two dummies for the ethnicity of the mother.[11]

We model survival status for four alternative time intervals: survival after 3, 6, 9, and 12 months. Consider, first, survival status after 3 months. Using control 1 infants as controls, we find that the grandmother effect is positive and significant. Inclusion of additional control variables does not change the results fundamentally, but reduces both the significance and the size of the effects (see columns 1

[9] Figures obtained from the Gender Info database from the United Nations Statistical Division.

[10] In the original DHS surveys, two geographical variables are included: a binary variable that distinguishes between rural an urban areas and a dichotomous variable that distinguishes between mountain, hill, or plain terrain. We create regional dummy variables obtained from the interaction of these two geographical variables.

[11] To ensure comparability between the surveys, we constructed an ethnicity variable that considers 6 ethnic groups. The ancestors of the *brahmin/chhetri* come from India. The *Newar* and the *Janajati*—who include many of Nepal's indigenous nationalities, such as the *Gurung*, the *Magar*, the *Tamang*, the *Tharu*, and the *Rai*— are sometimes referred to as old Nepalese groups. The *Muslin* are a minority in Nepal, comprising about 4% of the total population. The *Dalit*, sometimes referred to as "untouchables", are the lowest caste in the Hindu caste system. Finally, all the other ethnic groups, who represent around 10% of the population, are classified together. In the specifications, we report the results after controlling for a binary variable for Dalit, and a dummy variable for others. Using all the other dummies for ethnic categories does not change significantly the results and none of the other ethnic variables is significant in any of the specifications (results are available upon request).

Table 2: *Average Marginal Effects. Basic difference-in-differences results*

	3 months		**6 months**		**9 months**		**12 months**	
	Uncond.	*Cond.*	*Uncond.*	*Cond.*	*Uncond.*	*Cond.*	*Uncond.*	*Cond.*
Control 1	1	2	3	4	5	6	7	8
Female eligible	0.036	0.029	0.049	0.040	0.053	0.042	0.048	0.039
	(0.042)	(0.074)	(0.000)	(0.001)	(0.001)	(0.003)	(0.033)	(0.068)
Male eligible	-0.008	-0.015	-0.023	-0.036	-0.048	-0.064	-0.049	-0.091
	(0.873)	(0.777)	(0.700)	(0.572)	(0.502)	(0.413)	(0.489)	(0.307)
No. of obs.	7689	7424	7689	7424	7689	7424	7689	7424
Control 2								
Female eligible	0.056	0.042	0.068	0.049	0.067	0.049	0.072	0.053
	(0.011)	(0.129)	(0.000)	(0.030)	(0.000)	(0.034)	(0.001)	(0.031)
Male eligible	0.003	-0.028	-0.019	-0.078	-0.046	-0.112	-0.033	-0.126
	(0.965)	(0.743)	(0.795)	(0.459)	(0.577)	(0.330)	(0.666)	(0.289)
No. of obs.	1106	1015	1171	1079	1286	1183	1286	1206
Control 3								
Female eligible	0.033	0.030	0.046	0.041	0.048	0.042	0.042	0.038
	(0.114)	(0.085)	(0.002)	(0.001)	(0.003)	(0.003)	(0.107)	(0.094)
Male eligible	-0.016	-0.017	-0.032	-0.038	-0.058	-0.066	-0.062	-0.100
	(0.778)	(0.760)	(0.621)	(0.569)	(0.447)	(0.411)	(0.417)	(0.291)
No. of obs.	4369	4260	4443	4333	4503	4392	4503	4392
Control 4								
Female eligible	0.065	0.050	0.075	0.054	0.074	0.051	0.079	0.057
	(0.002)	(0.041)	(0.000)	(0.014)	(0.000)	(0.029)	(0.000)	(0.019)
Male eligible	0.000	-0.027	-0.029	-0.090	-0.052	-0.122	-0.038	-0.143
	(0.994)	(0.761)	(0.723)	(0.436)	(0.565)	(0.326)	(0.647)	(0.273)
No. of obs.	849	790	913	853	983	915	1003	933

Note: Average marginal effects as defined in equation (8). *p*-values are shown in parenthesis. "Months" refers to months after birth. Control 1 infants are infants who do not live with eligibles. Control 2 infants live with old people who are between 60 and 74 at the infant's birth date. Controls 3 infants live with people who are at most 60. Control 4 infants are infants who live with old people who are between 60 and 69. *Uncond.* refers to the diff-in-diff model with month of birth and region fixed effects. The *Cond.* model additionally includes dummy variables for the education of the mother, the mother's age at the child's birth and its square, a dummy for whether the child is female, the number of kids younger than five—at birth of infant—in the household, and dummies for the ethnicity of the mother.

and 2). In contrast to what we observe for the eligible female, the eligible male effect is not significant. The sign of the effect is, nevertheless, negative and larger after controlling for additional covariates.

Estimates for the determinants of survival rates after 6 months follow a similar pattern. Interestingly, the female eligible effect increases around 5 percentage points after 6 months in the unconditional model (first row in column 3) and 4 percentage points in the conditional model (first row columns 4). The point estimate for the male eligible effect is still negative and not significant. Similar results are obtained when looking at the survival determinants after 9 and 12 months. These results suggest that the positive effect on survival rates take place in the first 6 months after birth and only in the presence of a female eligible.

We can study the robustness of these results to alternative definitions of the control group. When we include as controls only infants who live in households where there are no old people—i.e. control 3 infants—the results are very similar to those using control 1 infants, arguably the result of the large demographic weight of these households. When we include as controls only those infants who live in households where there are old non-eligible people, the estimates of the female eligible effect become larger and are estimated more accurately. The negative effect of the male eligible effect, however, remains insignificant.

The largest point estimates of the female eligible effect are obtained when we use as controls the control 4 infants. Survival rates after 3 months improve around 6.5 percentage points according to the unconditional model and around 5 percentage points according to the conditional model. This effect increases after 6 months up to 7.5 percentage points and then it stabilizes to between 7.4 and 7.9 in the first year of the newborn. According to the conditional model results, these effects are somewhat smaller although still important: around 5.7 percentage points after a year.

4.2 Different effects for male and female infants

In Table 3 we report estimates of the female and male eligible effects for subsamples of only boys and only girls. For brevity, we only report the marginal effects using control 1 and 4 infants with the full set of additional covariates.[12]

Table 3: *Average Marginal Effects. Different effects for male and female infants*

	Control 1 infants				**Control 4 infants**			
	3 months	6 months	9 months	12 months	3 months	6 months	9 months	12 months
Only boys								
Female eligible	0.044	0.049	0.052	0.041	0.111	0.100	0.103	0.092
	(0.000)	(0.000)	(0.001)	(0.211)	(0.000)	(0.030)	(0.012)	(0.078)
Male eligible	-0.006	-0.058	-0.100	-0.150	-0.069	-0.290	-0.300	-0.329
	(0.913)	(0.527)	(0.372)	(0.252)	(0.705)	(0.247)	(0.214)	(0.159)
No. of obs.	3779	3779	3824	3824	315	336	364	390
Only girls								
Female eligible	0.018	0.036	0.035	0.038	0.084	0.080	0.080	0.146
	(0.624)	(0.039)	(0.123)	(0.128)	(0.024)	(0.030)	(0.018)	(0.000)
Male eligible	-0.055	-0.015	-0.007	-0.017	0.042	-0.275	-0.331	-0.382
	(0.135)	(0.598)	(0.771)	(0.572)	(0.482)	(0.442)	(0.311)	(0.203)
No. of obs.	3135	3257	3345	3423	152	171	184	202

Note: Average marginal effects as defined in equation (8) for subsamples of only boys and only girls. *p*-values are shown in parenthesis. "Months" refers to months after birth. Control 1 infants are infants who do not live with eligibles. Control 4 infants are infants who live with old people who are between 60 and 69. Controls are those in model *Cond.* defined in Table 2 and include month of birth and region fixed effects, dummy variables for the education of the mother, the mother's age at the child's birth and its square, a dummy for whether the child is female, the number of kids younger than five—at birth of infant—in the household, and dummies for the ethnicity of the mother.

The female eligible effect on boys is positive and generally significant—the only exception being the effect after 12 months using control 1 infants. Using control 4 infants, the estimate of the marginal effect on boys survival rates more than doubles. For example, after 12 months, survival rates improve by 4.1 percentage points using control 1 infants but they improve by 9.2 percentage points using as controls the more credible control 4 infants.

For girls, we also find a positive effect when there is a female eligible in the household. Using control 1 infants, the effect is smaller than for boys, and it is significant at the 10% level only after 6 months.

[12]Results using other controls and the unconditional model are similar. They are available upon request.

However, with control 4 infants the estimated marginal effects for girls are—in spite of the smaller sample—strongly significant. Interestingly, they become very close to those for boys (8 vs. 10 percentage points) after 3, 6, and 9 months and become larger after one year (14.6 vs. 9.2 percentage points). The male eligible effect is almost always negative. Perhaps more importantly, it is never significant, regardless of using the boys or the girls subsamples and both for control 1 and control 4 infants.[13]

Although we are agnostic as to the channels through which the effects take place, the findings are broadly in line with results reported elsewhere. In particular, the asymmetry between the male and female eligible effects replicate the basic results found in Duflo (2000, 2003). We do find that conditioning the sample by infant's gender does not alter neither the importance of the presence of a female eligible, nor the apparent absence of any effect in the presence of a male eligible. In contrast, in Duflo (2000) and Duflo (2003) the female eligible effect is significant only for girls. We observe a larger estimated effect for girls one year after birth.

We claim in Section 6 that this asymmetric results are driven by assuming Parallel Paths both for female–eligible and for male–eligible treatments. Before that, we need to rule out two alternative potential explanations. First, the absence of effects in the case of male eligibles in our sample could result from males strategically exploiting gender role differences in society. In that case, the male eligible effect would not be negative when he is the only eligible individual in the household. Second, the estimates presented so far rely on the assumption that the presence at the time of interview of an eligible person coincides with her or his presence at the time of the infant's birth. This is a strong assumption and could lead to spurious asymmetric results if female and male eligible individuals are not equally likely to remain in three-generation households.

4.3 Specialization

It could be argued that the absence of effects in the case of male eligibles masks a type of specialization pattern within the household. In the presence of a female eligible individual, the male eligible would expect the female to be the only contributor to the additional resources for the newborn. However, in the absence of differentials in gender preferences, if the male eligible is the only eligible individual in the household we would expect the male eligible effect to be of a similar magnitude to the female eligible effect.

We present in Table 4 separate estimates using as treated three alternative subgroups. In the first specification, we use only those households in which the male eligible is the only eligible individual in the household. In the second specification, we use only those households in which the female eligible is the only eligible individual in the household. If the effect estimated in the previous specifications is

[13] In the unconditional model, not reported in Table 3, the negative effects of the male eligible effect for girls are around -0.075 and significant at the 5% significance level when control 1 infants are used as controls.

just an artifact from gender specialization, then we would expect that the effects for the only–one–male eligible sample would be similar to the estimates reported until now for the female eligibles.

Table 4: *Gender specialization in the provision of household resources*

	Control 1 infants				Control 4 infants			
	3 months	6 months	9 months	12 months	3 months	6 months	9 months	12 months
Male eligible alone								
Male eligible		0.007	-0.036	-0.090		-0.022	-0.135	-0.203
		(0.913)	(0.721)	(0.512)		(0.870)	(0.529)	(0.403)
No. of obs.	7261	7290	7290	7290	585	666	720	748
No. of Treated	29	29	29	29	29	29	29	29
Female eligible alone								
Female eligible	0.016	0.024	0.023	0.002	0.004	0.009	-0.013	-0.024
	(0.601)	(0.391)	(0.508)	(0.965)	(0.960)	(0.900)	(0.892)	(0.816)
No. of obs.	7291	7291	7291	7291	642	711	768	785
No. of Treated	28	28	28	28	28	28	28	28
Neither is HH								
Female eligible	0.019	0.034	0.034	0.025	0.030	0.041	0.033	0.032
	(0.471)	(0.069)	(0.128)	(0.461)	(0.576)	(0.279)	(0.483)	(0.556)
Male eligible		-0.009	-0.006	-0.056		-0.012	0.000	-0.056
		(0.910)	(0.940)	(0.622)		(0.916)	(0.997)	(0.714)
No. of obs.	7321	7344	7344	7344	659	749	807	825
No. of Treated	57	57	57	57	57	57	57	57

Note: Average marginal effects as defined in equation (8) using only those households in which the male eligible is the only eligible individual in the household, using only those households in which the female eligible is the only eligible individual in the household, and using only those households when neither the male eligible nor the female eligible are reported as the household head. *p*-values are shown in parenthesis. "Months" refers to months after birth. Control 1 infants are infants who do not live with eligibles. Control 4 infants are infants who live with old people who are between 60 and 69. Controls are those in model *Cond.* defined in Table 2 and include month of birth and region fixed effects, dummy variables for the education of the mother, the mother's age at the child's birth and its square, a dummy for whether the child is female, the number of kids younger than five—at birth of infant—in the household, and dummies for the ethnicity of the mother. Absence of marginal effect estimates signals that a perfect prediction problem impedes the sample identification of the effect.

Due to the small treated sample, we encounter a perfect prediction problem when we try to estimate the effects for survival status after 3 months when the male eligible is alone. For all other models, the effects are usually negative for the male eligible sample and usually positive for the female eligible sample. However, due to the very small samples for the treated, the effects are never accurately estimated and we cannot reach any clear conclusion using this testing strategy.[14]

Alternatively, we look at the estimates of the effects when neither the male eligible nor the female eligible are reported as the household head. Nepal is a society where old people are frequently regarded as the most respected members of the family, even if they do not have a predominant economic position within the household. It is not rare to observe either a male or female eligible individual to be chosen as the household head, and we hypothesize that in many cases household head status is only a sign of respect to the elderly. We also assume that those elderly who are not chosen as household heads do not command a dominant economic position within the household.

If the specialization explanation is right, we would argue that in those households in which neither the

[14]Using a linear probability model we avoid the identification problems that we have with the probit model. For the only–one–male eligible sample, we find that the effects are negative and significant while in the only–one–female eligible sample the results are positive and significant. Results are available upon request.

male nor the female eligible are household heads, the differences between the effects of female and male eligible should be smaller. Turning to our results, we find that point estimates are again positive for the female eligible effect and negative for the male eligible. In the latter case, the effects are always non-significant. However, in the case of the female eligible effect, the effects are sometimes significant or borderline significant.

4.4 Retrospective information

The estimates presented so far rely on the assumption that the presence at the time of interview of an eligible person coincides with her or his presence at the time of the infant's birth. This assumption is less credible the larger the time span between the date of birth and the time of the interview. Therefore, in what follows we study the robustness of our results to alternative assumptions regarding retrospective information.

We consider in Table 5 three alternative samples. In the so-called *Minimized-delay sample*, we use for the pre-treatment period only infants born between June 1993 and June 1994. For the post-treatment period, we use infants born between June 1997 and June 1998. Defining our samples in this way, we make delay between births and the collection of information never larger than two and a half years for the pre-treatment period and never larger than three and a half years for the post-treatment period. Because of this asymmetry, we also consider the *Minimized-similar-delay sample* that includes all infants born between two and a half and three and a half years before the interview both for the pre-treatment period and for the post-treatment period. Consequently, in the *Minimized-similar-delay sample*, infants are born between June 1992 and June 1993 in the pre-treatment period and between June 1997 and June 1998 in the post-treatment period.

Both the *Minimized-* and the *Minimized-similar-delay* samples only include infants born within a period of 12 months. This greatly reduces the estimation sample and may potentially affect the accuracy of the estimates. Consequently, we additionally create a larger sample—that we refer to as *Similar-delay sample*—that includes children born between June 1991 and June 1993 for the pre-treatment period and children born between June 1996 and June 1998 for the post-treatment period. Births in this sample occur between two and a half and five years before the interview.

In Table 5 we present the basic diff-in-diffs estimates for the three alternative samples using the control 1 infants as the controls. When we consider the *Minimized-delay sample* we obtain positive point estimates for the female eligible effect in all specifications and negative point estimates for the male eligible effect in half of the specifications. However, all estimates are not significant with the exception of the male eligible effect 9 months after birth, which is positive and significant. Hence, changing the implicit assumption about retrospective information has an effect on the significance of our results.

The *Minimized-delay* pre-treatment sample differs in time span from the post-treatment sample. The

Table 5: *Retrospective Information: Results using Control 1 Infants*

	3 months		**6 months**		**9 months**		**12 months**	
	Uncond.	*Cond.*	*Uncond.*	*Cond.*	*Uncond.*	*Cond.*	*Uncond.*	*Cond.*
Minimize delay	1	2	3	4	5	6	7	8
Female eligible	0.011	0.014	0.021	0.021	0.021	0.022	0.001	0.008
	(0.804)	(0.626)	(0.565)	(0.394)	(0.607)	(0.451)	(0.982)	(0.860)
Male eligible	-0.047	-0.004	-0.066	-0.010	0.044	0.035	0.027	0.021
	(0.241)	(0.859)	(0.162)	(0.731)	(0.047)	(0.038)	(0.539)	(0.599)
No. of obs.	2657	2554	2657	2554	2674	2571	2674	2571
Minimize similar delay								
Female eligible	0.047	0.042	0.052	0.047	0.056	0.049	0.055	0.049
	(0.000)	(0.000)	(0.001)	(0.000)	(0.002)	(0.001)	(0.048)	(0.047)
Male eligible	-0.016	0.006	-0.013	-0.046	0.003	0.004	0.000	-0.053
	(0.556)	(0.774)	(0.615)	(0.383)	(0.971)	(0.953)	(0.996)	(0.670)
No. of obs.	2488	2403	2488	2403	2505	2420	2505	2420
Similar delay								
Female eligible	0.036	0.024	0.050	0.038	0.053	0.039	0.047	0.032
	(0.157)	(0.374)	(0.011)	(0.077)	(0.016)	(0.113)	(0.158)	(0.376)
Male eligible	-0.037	-0.045	-0.107	-0.126	-0.174	-0.191	-0.133	-0.253
	(0.710)	(0.648)	(0.441)	(0.383)	(0.292)	(0.259)	(0.301)	(0.171)
No. of obs.	5068	4891	5068	4891	5068	4891	5068	4891

Note: Average marginal effects as defined in equation (8) with alternative samples. The Minimized delay sample includes infants born between June 1993 and June 1994 for the pre-treatment period and infants born between June 1997 and June 1998 for the post-treatment period. The Minimized similar delay sample includes infants born between two and a half and three and a half years before the interview both for the pre-treatment period and for the post-treatment period. The Similar delay sample includes infants born between two and a half and five years before the interview. *p*-values are shown in parenthesis. "Months" refers to months after birth. Control 1 infants are infants who do not live with eligibles. *Uncond.* refers to the diff-in-diff model with month of birth and region fixed effects. The *Cond.* model additionally includes dummy variables for the education of the mother, the mother's age at the child's birth and its square, a dummy for whether the child is female, the number of kids younger than five—at birth of infant—in the household, and dummies for the ethnicity of the mother.

implicit assumption on retrospective information is less credible for the post-treatment period because there are many observations in that period for which the distance between the date of birth and the date of interview is larger than the largest distance in the pre-treatment period. This asymmetry could create a bias in our diff-in-diffs estimations. In contrast, using the *Minimized-similar-delay* sample we assume that there are no relevant changes in the composition of the families in the last three years prior to the interview, regardless of whether the interview takes places before or after the start of the implementation of the OAP programme. Interestingly, although the sample does not change much, the female eligible effect is now larger than the original estimates and even more significant while the male eligible effect is never significant.

One could nevertheless argue that the lack of significance for the positive point estimates of the male eligible effect after nine months (see columns 5 and 6) might be due to the small size of the estimation sample. To look at this possibility, we use the *Similar-delay* sample, which increases the sample size with respect to the *Minimize-similar-delay* sample but makes the same assumption on retrospective information both before and after the start of the policy. When we do this, the point estimates regarding the female eligible effect become very similar to the original results (although they are not significant in some cases). More interestingly, none of the point estimates of the male eligible effect are significant, and all of them are negative.

Assuming that the old people present in the household at the interview were already present at the date of the infant's birth is a potentially influential assumption. Our results show that alternative sample specifications (which imply alternative assumptions on retrospective information) lead to slight changes in the size of the effects and also to differences in the significance of the results. However, our results also suggest that the asymmetry between the female and the male eligible effects is not an artifact of this assumption. Moreover, in arguably the best alternative to our benchmark estimates—the estimates obtained from the *Minimized-similar-delay* sample—we find that the female eligible effect is positive and significant while the male eligible effect is not significant.

5 Testing common trends and an alternative identification strategy

So far, the identification of the effects of the OAP programme lies on the Parallel Paths assumption, which states that average changes in survival status among those treated if untreated are equal to the average changes in survival status among comparable controls. Violation of this assumption would lead to inconsistent estimates of the effect.

It is customary to test for common pre-treatment trends to justify the Parallel Paths assumption. The simplest way to do this is by conducting DID on the last pre-treatment period, a test that requires at least two periods before treatment. In our benchmark sample we include all children born between July 1991 and June 1994 for the pre-treatment period. We know the exact birth date for each observation so that for a sufficiently large sample we could consider as many periods as days in the pre-treatment period. However, since the number of treated is small even when we group them by month of birth, we opt for dividing the pre-treatment sample into only two periods: the first pre-treatment period includes all births from July 1991 to December 1992 while the second pre-treatment period includes all births between January 1993 and June 1994. We implement the test on pre-treatment common trends by using only the pre-treatment sample and then computing the diff-in-diffs estimator as if the policy had been implemented in January 1993 instead of July 1995. The test for common trends is the test on the significance of the diff-in-diffs estimate for the marginal effect α (see equation (8)).

Table 6 reports the results of the tests for pre-treatment common trends for the female eligible effect and the male eligible effect using all four alternative control groups and for both the unconditional and the conditional specifications.

We find no evidence of a female eligible effect in almost all specifications and periods considered. The only exceptions, for survival rates after 9 and 12 months, occur only in the unconditional model and using control 2 infants. With our preferred control group, control 4 infants, we cannot reject that controls and treated have common pre-treatment trends in survival status 3, 6, 9, and 12 months after birth.

Table 6: *Tests for pre-treatment common trends*

	3 months		6 months		9 months		12 months	
	Uncond.	*Cond.*	*Uncond.*	*Cond.*	*Uncond.*	*Cond.*	*Uncond.*	*Cond.*
Control 1	1	2	3	4	5	6	7	8
Female eligible	-0.001	-0.017	0.013	0.002	0.015	0.002	0.016	-0.001
	(0.984)	(0.811)	(0.796)	(0.969)	(0.781)	(0.968)	(0.784)	(0.990)
Male eligible	-0.952	-0.960	-0.946	-0.954	-0.939	-0.948	-0.933	-0.943
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
No. of obs.	3523	3365	3523	3365	3523	3365	3523	3397
Control 2								
Female eligible	0.033	0.023	0.037	0.029	0.043	0.032	0.046	0.034
	(0.356)	(0.537)	(0.140)	(0.237)	(0.037)	(0.169)	(0.032)	(0.129)
Male eligible	-0.975	-0.980	-0.977	-0.982	-0.974	-0.979	-0.973	-0.979
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
No. of obs.	503	473	558	523	610	571	610	571
Control 3								
Female eligible	-0.012	-0.036	0.005	-0.010	0.004	-0.010	0.005	-0.017
	(0.861)	(0.681)	(0.934)	(0.880)	(0.944)	(0.886)	(0.948)	(0.832)
Male eligible	-0.955	-0.962	-0.949	-0.957	-0.944	-0.951	-0.935	-0.945
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
No. of obs.	2064	1997	2064	1997	2064	1997	2064	2009
Control 4								
Female eligible	0.023	0.022	0.028	0.026	0.038	0.033	0.041	0.036
	(0.651)	(0.660)	(0.451)	(0.439)	(0.195)	(0.297)	(0.180)	(0.249)
Male eligible	-0.976	-0.978	-0.980	-0.982	-0.977	-0.979	-0.976	-0.978
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
No. of obs.	409	369	452	405	484	432	484	432

Note: Tests for pre-treatment common trends for the female eligible effect and the male eligible effect. *p*-values are shown in parenthesis. "Months" refers to months after birth. Control 1 infants are infants who do not live with eligibles. Control 2 infants live with old people who are between 60 and 74 at the infant's birth date. Controls 3 infants live with people who are at most 60. Control 4 infants are infants who live with old people who are between 60 and 69. *Uncond.* refers to the diff-in-diff model with month of birth and region fixed effects. The *Cond.* model additionally includes dummy variables for the education of the mother, the mother's age at the child's birth and its square, a dummy for whether the child is female, the number of kids younger than five—at birth of infant—in the household, and dummies for the ethnicity of the mother.

In contrast, we find very strong evidence of a negative and significant male eligible effect before treatment for all specifications, time ranges, and control groups. In the presence of pre-treatment trend differentials, Parallel Paths becomes less attractive as it implies that differing pre-treatment trends become equal after treatment under no treatment. Hence, these tests suggest that the results reported so far for the female eligible effect are based on a true assumption but the results for the male eligible effect are based on an assumption that is false.

What economic process could motivate different pre-treatment trends for male treated and controls, but the same pre-treatment trends for female treated and controls? One plausible explanation is the existence of trends in unobservable quality differentials by type of households. Suppose that couples live with elderly people if they cannot afford to live separately or if the elderly person needs their assistance because no-one else can help. As the economy develops and wages improve, the proportion of three-generation households where the young couple cannot afford to live separately will tend to decrease. When looking at successive cross-sections of data, this effect creates a downward trend in the relative average household wealth in three-generation households where the young couple cannot live separately. In contrast, economic growth does not improve the wealth of poor elderly who are not property owners. Hence, economic growth should not change the formation of three-generation households where the elderly requires assistance. As long as males are less affected by poverty than old females, old females will tend to live in three-generation households where the old person needs help while old males will live in households where the young couples are the ones who benefit economically from the association. When using successive cross-sections of three-generation households, economic progress will trigger a sample selection mechanism by which the average three-generation household with an old male will suffer a relative decline in wealth.

In practice, researchers who find pre-treatment trend differentials often formulate flexible econometric models to accommodate those trend differentials. In the next section, we follow Mora and Reggio (2012) and explore an alternative identification strategy using a flexible model and under an alternative assumption.

6 A flexible model with differing trends

Pre-treatment trend differentials in survival rates can be easily accommodated in the basic linear specification from equation (6) by including a time dummy for the last pre-treatment period, *LastPre*, and its interaction with the treated indicator D:

$$S^* = \beta_0 + \beta x + \gamma_D D + \gamma_L LastPre + \gamma_P Post + \phi_L D \times LastPre + \phi_P D \times Post + \varepsilon. \tag{9}$$

Under the Parallel Paths assumption, $\phi = \phi_P - \phi_L$. However, as argued in the previous subsection, the Parallel Paths assumption is not appealing for the male eligible treatment because of the presence of pre-treatment trend differentials between treated and controls. Mora and Reggio (2012) show that an alternative assumption which identifies the policy effect in the presence of pre-treatment trend differentials is the Parallel Growths assumption. Intuitively, Parallel Growths states that under no treatment the survival status of the treated would have experienced the same acceleration as the survival status of the controls. Assuming Parallel Growths leads to a difference-in-double-differences moment condition for ϕ:

$$\phi = E\left[\Delta^2 S^* | D = 1, x\right] - E\left[\Delta^2 S^* | D = 0, x\right]. \tag{10}$$

Hence, under equation (9) and Parallel Growths, $\phi = \phi_P - 2\phi_L$. The parameter of interest α can then be estimated using equation (8).

For the female eligible effect we do not expect a very different estimate under Parallel Growths than under Parallel Paths because our pre-treatment trend differentials tests suggest that, for the female eligible treatment, $\phi_L = 0$. In contrast, the results of our common trend tests suggest that, for the male eligible treatment, $\phi_L < 0$, and, hence, that $\phi_P - 2\phi_L > \phi_P - \phi_L$. We thus expect a larger estimate of the male eligible effect under Parallel Growths than under Parallel Paths. Intuitively, the Parallel Growths assumption implicitly takes into account that the treated infants living with male eligibles would have experienced a relative average decline in survival score under no treatment.

The added flexibility in equation (9) comes with a cost: in our data, sample identification of ϕ_P fails for the male eligible treatment under the most flexible trend specification and the benchmark sample. We present two strategies to overcome this problem: a) to extend the estimation sample to include observations from 1990; and b) to assume that, before 1993, infants face the same probability of survival regardless of the presence of a male eligible (i.e. $\gamma_D = 0$).

We provide support for the Parallel Growths assumption by testing that pre-treatment average acceleration was equal between the treated and the controls. As in the common-trends tests, we implement the test on pre-treatment common accelerations by using only the pre-treatment sample. We now partition the pre-treatment sample into three 12-month periods: from July 1991 to June 1992, from July 1992 to June 1993, and from July 1993 to June 1994. We then compute the diff-in-double-diffs estimator as if the policy had been implemented in July 1993 instead of July 1995. The test for common accelerations is the test on the significance of the diff-in-double-diffs estimate for the marginal effect α.

Table 7 reports the results of the tests for pre-treatment common accelerations for the female eligible effect and the male eligible effect using all four alternative control groups and for both the unconditional and the conditional specifications. When we tested common pretreatment trends, we found no evidence of a female eligible effect but very strong evidence of a negative and significant male eligible effect. In contrast, the results in Table 7 show that there is no evidence of differences in accelerations between treated and controls for the male eligible effects but the tests results suggest that the Parallel Growths assumption is not appropriate for the female eligible effect.

The results from Table 6 and Table 7 hint that the effect of the female eligible effect should be identified using the Parallel Paths assumption while the estimate of the male eligible effect should be identified using the Parallel Growths assumption. Hence, in Table 8 we report, using estimates of equation (9), estimated marginal effects under the Parallel Paths assumption for the female eligible effect and under the Parallel Growths assumption for the male eligible effect. For brevity, we only show the results using Control 4 infants.[15]

Table 8 corroborates the results so far concerning the female eligible treatment effect: results on survival status are always positive and significant for all time delays. Moreover, the size of the effects is very similar to the estimated effects assuming Parallel Paths and using both estimates from equation (6) and from equation (9).

The crucial novelty in Table 8 in relation to the results reported so far concerns the male eligible effect. Assuming Parallel Growths overturns the results obtained by using Parallel Paths: The male eligible

[15] All results are available upon request.

Table 7: *Tests for pre-treatment common accelerations*

	3 months		6 months		9 months		12 months	
	Uncond.	*Cond.*	*Uncond.*	*Cond.*	*Uncond.*	*Cond.*	*Uncond.*	*Cond.*
Control 1	1	2	3	4	5	6	7	8
Female Eligible	0.066	0.059	0.074	0.067	0.082	0.074	0.089	0.080
	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Male eligible	0.022	0.026	-0.042	-0.035	-0.072	-0.058	0.044	-0.061
	0.833	0.751	0.844	0.857	0.782	0.806	0.634	0.807
No. of obs.	4583	4382	4583	4382	4583	4382	4583	4428
Control 2								
Female Eligible	0.079		0.081	0.077	0.090	0.083	0.097	0.090
	0.000		0.000	0.000	0.000	0.000	0.000	0.000
Male eligible	0.064		0.017	0.026	-0.068	-0.040	0.049	-0.044
	0.128		0.907	0.838	0.811	0.873	0.606	0.866
No. of obs.	686		750	704	782	736	782	736
Control 3								
Female Eligible	0.067	0.060	0.076	0.068	0.084	0.077	0.094	0.085
	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Male eligible	0.014	0.012	-0.068	-0.088	-0.086	-0.094	0.041	-0.114
	0.913	0.919	0.789	0.746	0.763	0.743	0.702	0.716
No. of obs.	2728	2643	2728	2643	2728	2643	2728	2661
Control 4								
Female Eligible	0.086		0.091	0.084	0.096	0.087	0.102	0.093
	0.000		0.000	0.000	0.000	0.000	0.000	0.000
Male eligible	0.063		-0.006	0.001	0.011	0.020	0.079	0.022
	0.343		0.978	0.994	0.954	0.901	0.103	0.898
No. of obs.	531		581	550	616	582	616	582

Note: Tests for pre-treatment common accelerations for the female eligible effect and the male eligible effect. *p*-values are shown in parenthesis. "Months" refers to months after birth. Control 1 infants are infants who do not live with eligibles. Control 2 infants live with old people who are between 60 and 74 at the infant's birth date. Controls 3 infants live with people who are at most 60. Control 4 infants are infants who live with old people who are between 60 and 69. *Uncond.* refers to the diff-in-diff model with month of birth and region fixed effects. The *Cond.* model additionally includes dummy variables for the education of the mother, the mother's age at the child's birth and its square, a dummy for whether the child is female, the number of kids younger than five—at birth of infant—in the household, and dummies for the ethnicity of the mother.

effect changes from being negative and not-significant to being positive and strongly significant in most specifications. The only exceptions are estimates for the Only–boys sample after 6, 9, and 12 months in the conditional models where the point estimates are negative but very imprecisely estimated. These conclusions are similar to those obtained with alternative controls. Regarding, the size of the effect whenever is significant, it is closely similar to the size of the estimated effect for female eligible. In fact, we can never reject that the two effects are equal in size.

Finally, in the presence of different pre-treatment common trends it is usual in diff–in–diffs applications to extend the benchmark model for the male eligible effect by introducing a group–specific linear deterministic trend in equation (6). This is a more restrictive approach than identifying the male eligible effect using only the Parallel Growths assumption (Mora and Reggio, 2012) and this unnecessary restriction could bias the estimates. In results we do not show for brevity, we find that the estimates are also positive and significant but the point estimates are around 33% larger than the point estimates using only the Parallel Growths assumption. Hence, although the basic result remains (i.e., the male eligible effect is positive) we suspect that the deterministic linear trend specification introduces a positive bias.

Table 8: *Effects under Parallel Paths for female eligible and Parallel Growths for male eligible*

	3 months		**6 months**		**9 months**		**12 months**	
	Uncond.	*Cond.*	*Uncond.*	*Cond.*	*Uncond.*	*Cond.*	*Uncond.*	*Cond.*
				Pre-treatment sample: July 1990-June 1994				
All								
Female effect	0.063	0.053	0.074	0.059	0.071	0.056	0.075	0.062
	0.006	0.030	0.000	0.009	0.006	0.035	0.011	0.026
Male effect	0.072	0.062	0.077	0.065	0.078	0.064	0.086	0.070
	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Difference (p-value)	0.450	0.403	0.696	0.621	0.658	0.628	0.403	0.587
No. of obs.	978	916	1042	979	1115	1044	1135	1062
Only boys								
Female effect	0.112	0.107	0.117	0.109	0.116	0.111	0.094	0.101
	0.000	0.000	0.000	0.002	0.004	0.001	0.164	0.019
Male effect	0.105	0.094	0.066	-0.018	0.069	-0.024	0.113	-0.013
	0.000	0.000	0.608	0.947	0.580	0.932	0.014	0.962
Difference (p-value)	0.837	0.918	0.617	0.446	0.697	0.421	0.770	0.522
No. of obs.	426	379	450	400	487	430	508	456
				Model with $\gamma_D = 0$				
All								
Female effect	0.064	0.051	0.073	0.054	0.069	0.050	0.072	0.056
	0.008	0.051	0.000	0.022	0.014	0.072	0.019	0.051
Male effect	0.062	0.052	0.062	0.045	0.058	0.036	0.075	0.039
	0.023	0.065	0.051	0.267	0.179	0.504	0.003	0.527
Difference (p-value)	0.939	0.913	0.899	0.901	0.918	0.851	0.810	0.821
No. of obs.	849	790	913	853	983	915	1003	933
Only boys								
Female effect	0.118	0.114	0.119	0.108	0.115	0.110	0.092	0.102
	0.000	0.000	0.002	0.008	0.018	0.003	0.237	0.027
Male effect	0.096	0.067	0.028	-0.127	0.022	-0.140	0.091	-0.167
	0.032	0.502	0.871	0.730	0.902	0.708	0.273	0.672
Difference (p-value)	0.840	0.562	0.497	0.297	0.546	0.268	0.948	0.289
No. of obs.	359	315	383	336	418	364	439	390

Note: Average marginal effects obtained using equation (9) and assuming Parallel Paths for the female eligible and Parallel Growths for the male eligible effect. Panel Pre-treatment sample: July 1990-June 1994 extends the estimation sample to include observations from 1990. Panel Model with $\gamma_D = 0$ restricts γ_D in equation (9). p-values are shown in parenthesis. "Months" refers to months after birth. Control 4 infants are infants who live with old people who are between 60 and 69. *Uncond.* refers to the diff-in-diff model with month of birth and region fixed effects. *Uncond.* refers to the diff-in-diff model with month of birth and region fixed effects. The *Cond.* model additionally includes dummy variables for the education of the mother, the mother's age at the child's birth and its square, a dummy for whether the child is female, the number of kids younger than five—at birth of infant—in the household, and dummies for the ethnicity of the mother.

To sum up, our results show that the Parallel Path assumption is essential to find differences between grandmothers and grandfathers. Under the Parallel Growths assumption and a flexible specification for the econometric model, we find no gender differences in how a positive shock in income among old people affects the welfare of infants living with them.

7 Conclusions

Many studies find evidence that presence of a grandmother is associated with higher child survival rates while no such association is found in the case of a grandfather. We exploit income variation from the introduction of a non-contributory universal pension scheme in Nepal in 1995. Using cross-sectional data from the 1996 and 2001 Nepal Demographic and Health Surveys, we obtain diff–in-diffs estimates that are consistent with these results: we find positive and significant effects on survival rates for an income increase of a female person older than 75 who lives in the same household while negative and sometimes significant effects for the income increase of an old male.

These results are qualitatively similar across alternative definitions of the control group, for both boys and girls, and do not depend on: a) how we exploit retrospective information in the data; b) whether the female (male) eligible is the only eligible in the household; or c) the family status of the eligible. However, the results are not robust to alternative assumptions for the diff–in–diffs estimates. More precisely, when we implement a flexible identification strategy based on the Parallel Growths assumption defined in Mora and Reggio (2012) for the male eligible effect and on the Parallel Paths assumption for the female eligible effect, we find no significant gender differences in how grandparents' economic conditions affect infant survival rates.

We validate the Parallel Growths assumption with a pre-treatment common acceleration test that is similar in spirit to the pre-treatment common tests used to validate the Parallel Paths assumption. We motivate the different results of these tests for female and male beneficiaries by the following argument. If couples tend to live with a male beneficiary when they have economic problems and tend to live with a female beneficiary when she has economic problems, then economic growth will result in successive cross-sections where three-generation households with an old male will suffer a relative decline in wealth. Hence, our findings can be interpreted as suggestive that cross-sectional analysis may bias downwards the estimates of the effect of grandfathers.

We leave for future work the study of the channels through which the effects reported in this paper take place. In particular, it is not clear whether the effects are just a consequence of higher income in the treated households. For example, if risk averse parents respond to the program with higher fertility, then the results found would be compatible with these type of parents having better success in bringing up kids. In addition, if only those eligible that worry more about the family end up collecting the benefits, then our results would likely underestimate the average effect of the income effect.

References

CASTERLINE, J. B. (1989): "Collecting data on pregnancy loss: a review of evidence from the World Fertility Survey," *Studies in Family planning*, 20(2), 81–95.

DUFLO, E. (2000): "Child Health and Household Resources in South Africa: Evidence from the Old Age Pension Program," *The American Economic Review*, 90(2), pp. 393–398.

——— (2003): "Grandmothers and Granddaughters: Old-Age Pensions and Intrahousehold Allocation in South Africa," *The World Bank Economic Review*, 17(1), 1–25.

MORA, R., AND I. REGGIO (2012): "Treatment effect identification using alternative parallel assumptions," Discussion paper.

RAJAN, S. I. (2003): "Oldage Allowance Program in Nepal," http://www.chronicpoverty.org/uploads/publication_files/CP_2003_Rajan.pdf, [Online; accessed 11-June-2013].

SEAR, R., AND R. MACE (2008): "Who keeps children alive? A review of the effects of kin on child survival," *Evolution and Human Behavior*, 29(1), 1 – 18.

SHARMA, K. (2006): "The political economy of civil war in Nepal," *World Development*, 34(7), 1237 – 1253.

THAPA, S. (2004): "Abortion law in Nepal: the road to reform," *Reproductive Health Matters*, 12(24), 85–94.

THAPA, S., AND S. M. PADHYE (2001): "Induced abortion in urban Nepal," *International Family Planning Perspectives*, pp. 144–151.

Chapter 3: The Effect on Citation Inequality of Differences in Citation Practices across Scientific Fields

1. INTRODUCTION

The field dependence of reference and citation counts in scientific articles has been recognized since the beginning of Scientometrics as a field of study (see *inter alia* Pinski and Narin, 1976, Murugesan and Moravcsik, 1978, and Garfield, 1979). There are multiple reasons. Consider the differences across scientific disciplines in, for example, (i) size, measured by the number of publications in the periodical literature; (ii) the average number of authors per paper; (iii) the average paper length; (iv) the average number of papers per author in a given period of time; (v) the theoretical or experimental mix that characterizes each discipline; (vi) the average number of references per paper; (vii) the proportion of references that are made to other articles in the periodical literature; (viii) the percentage of internationally co-authored papers, or (ix) the speed at which the citation process evolves.

This paper develops a measuring framework where it is possible to quantify the importance of differences in citation practices. We use a model in which the number of citations received by an article is a function of two variables: the article's underlying scientific influence, and the field to which it belongs. In this context, the citation inequality of the distribution consisting of all articles in all fields –the all-fields case– is the result of two forces: differences in scientific influence, and differences in citation practices across fields. The first aim of the paper is how to isolate the citation inequality attributable to the latter, and how to measure its importance relative to overall citation inequality of all sorts.

The first difficulty we must confront is that the characteristics of the scientific influence distributions are *a priori* unknown. Thus, even if they were observable, we would not know how to compare the scientific influence of any two articles belonging to different fields. To overcome this difficulty, we make the strong assumption that articles in the same quantile of the scientific influence distribution have the same *degree* of scientific influence independently of the field to which they belong. Thus, if your article and mine belong, for example, to the 80^{th} percentile of our respective distributions, then we assume that they have the same degree of scientific influence.

The next difficulty is that scientific influence is an unobservable variable. To overcome this difficulty, we may remain agnostic about the myriad of motives researchers have in their citation behavior as long as we are allowed to assume that citation impact varies monotonically with scientific influence (for a survey of the controversies concerning the meaning of citation counts, see Bornmann and Daniel, 2008). Thus, if one article has greater scientific influence than another one in the same homogeneous field, then we expect the former to have also a greater citation impact than the latter.[1] The monotonicity assumption ensures that, for any field, the quantiles of the (unobservable) scientific influence distribution coincide with the quantiles of the corresponding (observable) citation distribution. Therefore, if the mean citation of articles in the 80^{th} percentile of your field is, for example, twice as large as the mean citation of articles in the same percentile in my field, this means

[1] The idea that citations is an observable indicator for a latent concept of scientific or scholarly influence, as well as the monotonicity assumption, are also found in Ravallion and Wagstaff (2011) in a different scenario: the construction of bibliometric measures of research impact.

that your field uses double number of citations than mine to represent the same status in scientific influence. The implication is that the citation inequality observed at any quantile can be solely attributed to idiosyncratic differences in citation practices. Thus, the aggregation of this measure over all quantiles provides a method of quantifying the effect of these differences (This is, essentially, John Roemer's, 1998, model for the study of inequality of opportunities in an economic or sociological context).

We implement this model by using an additively decomposable inequality index, in which case the citation inequality attributed to differences in citation practices is captured by a between-group inequality term in the double partition by field and citation quantile (Ruiz-Castillo, 2003). Specifically, using a dataset of 4.4 million articles published in 1998-2003 with a five-year citation window and an appropriate citation inequality index, we estimate that the citation inequality attributable to differences in citation practices across the 22 fields (219 sub-fields) distinguished by Thomson Scientific represents about 14% (18%) of overall citation inequality.

It would appear that, regardless of how their impact can be measured, differences in publication and citation practices pose insurmountable obstacles to direct comparisons of the absolute number of citations received by articles in different fields. For example, in the dataset used in this paper, how can we interpret the fact that the mean citation in Mathematics is 2.4, about eight and a half times smaller than in Molecular Biology and Genetics where it is equal to 20.4 citations? This paper shows that the striking similarity between citation distributions (documented at different aggregation levels in Albarrán and Ruiz-Castillo, 2011, Albarrán *et al.*, 2011, and Radicchi and Castellano, 2012a), causes the citation inequality attributable to different citation practices to be approximately constant over a wide range of quantiles. This allows its effect to be rather well estimated over that interval. Consequently, we provide a set of *exchange rates* and their standard deviations (StDevs hereafter) that serve to answer the following two questions. Firstly, how many citations in a given field are equivalent to, say, 10 citations in the all-fields case? For example, in *Clinical Medicine* the answer is 12.1 with a StDev of 0.6, while in *Mathematics* the answer is 3.3 with a StDev of 0.2. Secondly, how much can we reduce the effect of different citation practices by normalizing the raw citation data with the exchange rates? We find that this normalization procedure reduces this effect from 14% (18%) to around 2% (3.8%) of overall citation inequality.

The difficulty of comparing citation counts across scientific fields is a very well known issue that has worried practitioners of Scientometrics since its inception. Differences in citation practices are usually taken into account by choosing the world mean citation rates as normalization factors (see *inter alia* Moed *et al.*, 1985, 1988, 1995, Braun *et al.*, 1985, Schubert *et al.*, 1983, 1987, 1988, Schubert and Braun, 1986, 1996, and Vinkler 1986, 2003). More recently, other contributions support this traditional procedure on different grounds (Radicchi *et al.*, 2008, Radicchi and Castellano, 2012a, 2012b). In our last contribution, we find that using field mean citations as normalization factors leads practically to the same reduction of the effect of differences in citation practices on citation inequality as our exchange rates. We show how our model helps explaining why the traditional model is so successful.[2]

---

[2] Methods that use mean citations or exchange rates as normalization factors belong to the class of target or "cited side" normalization procedures. Following an idea in Small and Sweeney (1985), source or "citing side" procedures have been

The rest of the paper consists of five Sections. Section II is devoted to a review of the literature on normalization using field citation means. Section III introduces the model for the measurement of the effect of differences in citation practices, while Section IV contains an estimate of this effect in term of an appropriate additively decomposable citation inequality index. Section V presents the estimation of average-based exchange rates and its StDevs over a large quantile interval, and discusses the consequences of using such field exchange rates and mean citations as normalization factors. Section VI contains some concluding comments.

2. A REVIEW OF THE LITERATURE

From an operational point of view, a scientific field is a collection of papers published in a set of closely related professional journals. A field is said to be *homogeneous* if the number of citations received by its papers is comparable independently of the journal where each has been published. The problem we confront in this paper arises when one wants to evaluate research units publishing in closely related but heterogeneous fields –such as a Chemistry department working in Organic and Inorganic Chemistry– or, more simply, when one wants to directly compare the citations received by two papers in different scientific fields at any aggregation level.

As indicated in the Introduction, the traditional solution is to rely on the world mean citation in each field as the normalization factor. Note that no confidence interval is usually provided in applications of this normalization procedure. This is probably due to the high variances that characterize highly skewed citation distributions (for the 22 fields covered in this paper, see column 4 in Table A1 in the Appendix). More importantly, no deep explanation is usually given for mean normalization. It is simply agreed that the field mean citation captures well the expected value with which actual citation counts in that field can be related in order to compare normalized ratios across fields.

Let s_1 be the mean of a citation distribution, and let s_2 be the mean of those articles with citations above s_1. Under the idea that the difference $(s_2 – s_1)$ is a very good proxy for the StDev of citation distributions, Glänzel (2011) suggests a normalization of the raw data using this average-based difference.

In an important move, Radicchi *et al.* (2008) and Radicchi and Castellano (2012b) have recently justified the traditional solution on strong empirical grounds, namely, the *universality claim* according to which citation distributions in all fields exclusively differ by a scale factor. However, using a large dataset of 3.7 million articles published in 1998-2002, Albarrán *et al.* (2011) establish that the universality claim fails at both ends of the citation distributions at different aggregation levels, including a set of 219 sub-fields identified with the Web of Science subject-categories distinguished by Thomson Scientific (using a different methodology, Waltman *et al.*, 2011 reach the same conclusion). In the first place, Albarrán *et al.* (2011) find that the existence of a power law cannot be rejected at the top of the upper tail in 140 out of 219 sub-fields. On average, power laws represent 2% of all articles

recently suggested (see *inter alia* Zitt and Small, 2008, Moed, 2010, and Leydesdorff and Opthof, 2010). Since our dataset lacks citing side information, applying this type of procedure is beyond the scope of this paper.

in a sub-field, and account for about 13.5% of all citations. However, the large dispersion of the power law parameters is a clear indication that excellence is not equally structured in all citation distributions.[3] In the second place, the proportion of articles without citations and with some citations below the mean at the sub-field level represent on average 24.7% and 43.9% of all articles, respectively, with large SDs equal to 13.9 and 12.5. Possibly, this is partly due to the fact that a common five-year citation window was taken for all sub-fields in spite of the large differences in the time that it takes for citation processes to reach a given degree of completion.

This assessment contrasts with the more optimistic view in Radicchi *et al.* (2008) that supports the universality claim with a methodology that does not inform about how to treat the assignment of articles to multiple sub-fields, omits articles without citations, examines distributions at a limited set of points and, above all, covers only 14 of the 219 sub-fields. Radicchi and Castellano (2012b), which is free from other methodological shortcomings, focus only on 10 sub-fields within Physics. However, in a very important and more recent contribution that uses a dataset of about three million papers, covering 172 subject-categories, Radicchi and Castellano (2012a) –RC hereafter– also reject the universality claim. This seems to preclude certain normalization procedures. *"Making citation counts independent of the subject-categories seems therefore not possible with the use of linear transformations, because the difference between citation distributions of different subject-categories is not only due to a single scaling factor."* (RC, p. 2). More generally, *"A universal criterion for the complete suppression of differences among scientific domains probably does not exist. There are too many factors to account for, and consequently the 'philosophy' at the basis of a 'fair' normalization procedure is subjective"* (RC, p. 7). Nevertheless, RC demonstrate that, provided one is prepared to make strong assumptions, it is possible to find interesting normalization procedures. Ultimately, these normalization procedures work well in practice due to the similarity between citation distributions –a crucial aspect that deserves a few lines.

Generally, citation distributions are very different in many respects and, particularly, in size and mean citation rates. Consequently, it is very useful to use a size- and scale-invariant approach in order to focus on the shape of such distributions. One example is the Characteristic Scores and Scales (CSS hereafter) technique, introduced by Schubert *et al.* (1987) in the analysis of citation distributions. The CCS permits the partition of any citation distribution into a number of classes as a function of their members' citation characteristics. The following *characteristic scores* are determined: s_1 = mean of a citation distribution; s_2 = mean citation of articles with citations above s_1, and s_3 = mean citation of articles with citations above s_2. Although there is no universal distribution over the entire domain of all fields at any aggregate level, striking similarities over a broad partition of citation distributions at all aggregate levels have been found. In particular, on average, the proportion of articles at different aggregation levels that (i) receive none or few citations below s_1, (ii) are fairly well cited, namely, with citations between s_1 and s_2, and (iii) are remarkably or outstandingly cited with citations above s_2 is, approximately, 69/21/10. These three classes of articles account for the proportions 21/34/45 of all

[3] In addition, consider the possibility of defining a high-impact indicator over the sub-set of articles with citations above the 80th percentile of citation distributions. The distribution of high-impact values for the 219 sub-fields according to an indicator of this type is highly skewed to the right, and it presents some important extreme observations (see Herranz and Ruiz-Castillo, 2012).

citations. The small StDevs that come with these average values establish the strong similarity between such highly skewed citation distributions (see Table 6 in Albarrán *et al.*, 2011a, and Figure 2 in Albarrán and Ruiz-Castillo, 2011).[4]

In this scenario, RC's scheme is based on the assumption that *"each discipline or field of research has the same importance for the development of scientific knowledge. A fair numerical indicator, based on citation numbers, must then assume values that do not depend on the particular scientific domain under consideration. Under this assumption, the probability to find a paper with a given value of the fair indicator must not depend on the discipline of the paper, or equivalently, the distribution of normalized citation counts must be the same for all disciplines."* (RC p. 7). RC's main result is that the transformation of raw citation numbers that makes the normalized citation distributions the same for all fields is a non-linear function that depends on only two parameters for every field: the mean, and an exponential factor that are rather stable over different publication years from 1980 to 2004. Moreover, mirroring the similarities between citation distributions just documented, RC find strong regularities: the exponential factor assumes approximately the same value for the vast majority of 172 subject-categories, suggesting that –after all– the main difference between the citation distributions of different subject-categories is given only by a scale factor. Consequently, the rescaling advocated in Radicchi *et al.* (2008) and Radicchi and Castellano (2012a) using simply the mean, despite not being strictly correct, seems a very good approximation of the transformation able to make citation counts not depending on the scientific domain.

3. THE MODEL

3. 1. Notation and Assumptions

Let N_f be the total number of articles in a homogeneous field f, and let $\boldsymbol{c_f} = (c_{f1},\ldots,c_{fN_f})$ be the citation distribution for that field where, for each $i = 1,\ldots, N_f$, c_{fi} is the number of citations received by the i-th article. Assume that there are F homogeneous fields, indexed by $f = 1,\ldots, F$. The total number of articles in the all-fields case is $N = \Sigma_f N_f$. The number of citations of any article, c_{fi}, is assumed to be a function of two variables: the field f to which the article belongs, and the scientific influence of the article in question, q_{fi}, which is assumed for simplicity to be a single-dimensional variable. Thus, for every f we write:

$$c_{fi} = \phi(f, q_{fi}),\ i = 1,\ldots, N_f. \tag{1}$$

Let $\boldsymbol{q_f} = (q_{f1}, q_{f2},\ldots, q_{fN_f})$ with $q_{f1} \leq q_{f2} \leq\ldots\leq q_{fN_f}$ be the ordered distribution of scientific influence in every field. It is important to emphasize that distribution $\boldsymbol{q}_f$ is assumed to be a characteristic of the field. Furthermore, no restriction is *a priori* imposed on distributions $\boldsymbol{q_f}$, $f = 1,\ldots, F$. Consequently, for any two articles i and j in two different fields f and g, the values q_{fi} and q_{gj} cannot be directly compared. To overcome this difficulty, in this paper we introduce some structure into the comparability problem by means of the following key assumption.

[4] We saw before that articles receiving either no or few citations at the sub-field level had large StDevs. However, because of a strong negative correlation between these two groups, the broad class of poorly cited articles with citations below s_1 is located on average –as we have seen– around the 69th percentile of citation distributions with a StDev of 3.7.

Assumption 1a (A1a). *Articles at the same quantile π of any field scientific influence distribution have the same degree of scientific influence in their respective field.*

Typically, scientific influence is an unobservable variable. However, although the form of ϕ in Eq. 1 is unknown, we adopt the following assumption about it:

Assumption 2a (A2a). *The function ϕ in expression (1) is assumed to be monotonic in scientific influence, that is, for every pair of articles i and j in field f, if $q_{fi} \leq q_{fj}$, then $c_{fi} \leq c_{fj}$.*

Under A2a, the degree of scientific influence uniquely determines the location of an article in its field citation distribution. In other words, for every f, the partition of the scientific influence distribution $\boldsymbol{q_f}$ into Π quantiles of size N_f/Π, $\boldsymbol{q_f} = (\boldsymbol{q_f^1}, \ldots, \boldsymbol{q_f^\pi}, \ldots, \boldsymbol{q_f^\Pi})$, induces a corresponding partition of the citation distribution $\boldsymbol{c_f} = (\boldsymbol{c_f^1}, \ldots, \boldsymbol{c_f^\pi}, \ldots, \boldsymbol{c_f^\Pi})$ into Π quantiles, where $\boldsymbol{c_f^\pi}$ is the vector of the citations received by the N_f/Π articles in the π-th quantile of field f. Assume for a moment that we disregard the citation inequality within every vector $\boldsymbol{c_f^\pi}$ by assigning to every article in that vector the mean citation of the vector itself, namely, μ_f^π. Since the quantiles of citation impact correspond –as we have already seen– to quantiles of the underlying scientific influence distribution, holding constant the degree of scientific influence at any level as in A1 is equivalent to holding constant the degree of citation impact at that level. Thus, the interpretation of the fact that, for example, $\mu_f^\pi = 2\,\mu_g^\pi$ is that, on average, field f uses twice the number of citations as field g to represent the same underlying phenomenon, namely, the same degree of scientific influence in both fields. Hence, for any π, the difference between μ_f^π and μ_g^π for articles with the same degree of scientific influence is entirely attributable to differences in citation practices between the two fields.

Welfare economists would surely recognize the above as Roemer's (1998) model for the inequality of opportunities where individual incomes (or other indicators of performance, such as educational outcomes) are assumed to be a function of two types of factors: a set of variables outside an individual's responsibility –the *circumstances*, mainly inherited from our parents–, and *effort*, an unobservable single dimensional variable entirely within the sphere of each individual's responsibility. Circumstances allow a partition of the population into *types*. The distribution of effort within each type is assumed to be a characteristic of the type. Consequently, the amounts of effort exercised by individuals from different types are not ethically comparable. However, degrees of effort, measured by quantiles of the effort distribution for each type, are assumed to be comparable (A2a). Under the monotonicity assumption (A1a), quantiles of effort are seen to correspond with observable income quantiles. In this model, income inequality holding constant the degree of effort by every type is seen to be entirely due to differences in circumstances, or to the *inequality of opportunities* at this degree of effort. Income inequality due to differences in effort is not worrisome from a social point of view. It is income inequality due to differences in circumstances, namely, the inequality of opportunities, what society might attempt to compensate for. Individuals are articles, the equivalent of income is citations, types are fields, and effort is scientific influence.

3.2. The Measurement of the Effect of Differences in Citation Practices

For any population partition, we are interested in expressing the overall citation inequality as the sum of two terms: a weighted sum of *within-group* inequalities, plus a *between-group* inequality component. An inequality index is said to be *decomposable by population subgroup*, if the decomposition procedure of overall inequality into a within-group and a between-group term is valid for any arbitrary population partition. In the relative, or scale-invariant inequality case it is customary to calculate the between-group component by applying the inequality index to a citation vector in which each article in a given subgroup is assigned the subgroup's citation mean. Under this convention, it is well known that the Generalized Entropy (GE hereafter) family of inequality indices are the only measures of relative inequality that satisfy the usual properties required from any inequality index[5] and, in addition, are decomposable by population subgroup (Bourguignon, 1978, and Shorrocks, 1980, 1984).

Let $\boldsymbol{C} = (c_1, \ldots, c_l, \ldots, c_N)$ be the overall citation distribution in the all-fields case, where, for each l, there exists some article i in some field f such that $c_l = c_{fi}$. Without loss of generality, it is useful to develop the following measurement framework in terms of only one member of this family, the first Theil index, denoted by I_1, and defined as:

$$I_1(\boldsymbol{C}) = (1/N)\ \Sigma_l\ (c_l/\mu)\ \log\ (c_l/\mu), \tag{2}$$

where μ is the mean of distribution $\boldsymbol{C}$. The formula for the I_1 index when written in decomposable form for the partition of $\boldsymbol{C}$ into Π quantiles, $\boldsymbol{C} = (c^1, \ldots, c^\pi, \ldots, c^\Pi)$, is the following

$$I_1(\boldsymbol{C}) = \Sigma_\pi v^\pi I_1(\boldsymbol{c}^\pi) + I_1(\boldsymbol{\mu}^1, \ldots, \boldsymbol{\mu}^\Pi), \tag{3}$$

where v^π is the share of total citations received by articles in quantile π, and $I_1(\boldsymbol{\mu}^1, \ldots, \boldsymbol{\mu}^\Pi)$ is the citation inequality of the distribution $\boldsymbol{m} = (\boldsymbol{\mu}^1, \ldots, \boldsymbol{\mu}^\Pi)$ in which each article in a given quantile π is assigned the quantile's citation mean, $\mu^\pi = \Sigma_f [(N_f/N]\mu_f^\pi$. Next, for each π, the decomposability property of I_1 is applied to the partition into F fields, $\boldsymbol{c}^\pi = (\boldsymbol{c}_1^\pi, \ldots, \boldsymbol{c}_F^\pi)$:

$$I_1(\boldsymbol{c}^\pi) = \Sigma_f v_f^\pi I_1(\boldsymbol{c}_f^\pi) + I_1(\boldsymbol{\mu}_1^\pi, \ldots, \boldsymbol{\mu}_F^\pi), \tag{4}$$

where v_f^π is the share of total citations in quantile π received by articles in $\boldsymbol{c}_f^\pi$, and $I_1(\boldsymbol{\mu}_1^\pi, \ldots, \boldsymbol{\mu}_F^\pi)$ is the citation inequality of the distribution in which each article in quantile π of field f receives that sub-group's mean citation $\boldsymbol{\mu}_f^\pi$. Inserting (4) into (3), overall citation inequality is seen to be:

$$I_1(\boldsymbol{C}) = W + S + IDCP, \tag{5}$$

where:

$$W = \Sigma_\pi v^\pi \Sigma_f v_f^\pi I_1(\boldsymbol{c}_f^\pi) = \Sigma_\pi \Sigma_f v^{\pi,f} I_1(\boldsymbol{c}_f^\pi)$$

$$S = I_1(\boldsymbol{\mu}^1, \ldots, \boldsymbol{\mu}^\Pi)$$

---

[5] Namely, continuity; scale invariance; invariance to population replications, or size-invariance, and S-convexity that ensures that transfers from an article with more citations to another with fewer citations without altering their ranking reduces citation inequality.

$$IDCP = \Sigma_{\pi} v^{\pi} I_1(\boldsymbol{\mu}_1^{\pi},\ldots, \boldsymbol{\mu}_F^{\pi}),$$

where $v^{\pi,f}$ is the share of total citations received by articles in $\boldsymbol{c}_f^{\pi}$. The term W in equation (5) is a within-group term, which captures the weighted citation inequality within each quantile in every field. Obviously, since all articles in each vector $\boldsymbol{c}_f^{\pi}$ belong to the same field, there is no difficulty in computing the expression $I_1(\boldsymbol{c}_f^{\pi})$. Clearly, for large Π, $I_1(\boldsymbol{c}_f^{\pi})$, and hence term W is expected to be small. The term S is the citation inequality of the distribution $\boldsymbol{m} = (\boldsymbol{\mu}^1,..., \boldsymbol{\mu}^{\Pi})$ in which each article in a given quantile π is assigned the quantile's citation mean, $\mu^{\pi} = \Sigma_f [(N_f/N]\mu_f^{\pi}$. Thus, S is a measure of citation inequality at different degrees of citation impact that captures well the skewness of science in the all-fields case. Due to the high skewness of all citation distributions (see *inter alia* Albarrán and Ruiz-Castillo, 2011, and Albarrán *et al.*, 2011), the term S is expected to be large. Finally, for any π, the expression $I_1(\boldsymbol{\mu}_1^{\pi},\ldots, \boldsymbol{\mu}_F^{\pi})$, abbreviated as $I(\pi)$, is the citation inequality attributable to differences in citation practices according to I_1. Thus, the weighted average that constitutes the third term in Eq. 5, denoted by *IDCP* (*I*nequality due to *D*ifferences in *C*itation *P*ractices), provides a good measure of the citation inequality due to such differences.

4. THE ESTIMATION OF THE EFFECT OF DIFFERENCES IN CITATION PRACTICES

4. 1. The Data

Since we wish to address a homogeneous population, in this paper only research articles or, simply, articles, are studied. The dataset consists of 4.4 million articles published in 1998-2003, and the 35 million citations they receive after a common five-year citation window for every year, namely, citations received from 1998 to 2002 for articles published in 1998, up to 2003 to 2007 for articles published in 2003.

We study two field classifications in this paper, i.e. 22 broad fields and 219 sub-fields distinguished by Thomson Scientific. Table A1 and Table A2 in the Appendix presents the number of articles and mean citation rates for the 22 field and 219 sub-field cases respectively.

4. 2. The Choice of Inequality Index and the Number of Quantiles

The GE family can be described by means of the following convenient cardinalization:

$$Ia(\boldsymbol{C}) = (1/N)\ (1/\ a^2 - a)\ \Sigma_l\ (c_l/\mu^a - 1),\ a \neq 0,1; \tag{6}$$

$$I_0(\boldsymbol{C}) = (1/N)\ \Sigma_l \log (\mu/c_l);$$

$$I_1(\boldsymbol{C}) = (1/N)\ \Sigma_l\ (c_l/\mu) \log (c_l/\mu).$$

Parameter a summarizes the sensitivity of I_a in different parts of the productivity distribution: the more positive (negative) a is, the more sensitive I_a is to differences at the top (bottom) of the distribution (Cowell and Kuga, 1981). I_1 is the original Theil index, while I_0 is the mean logarithmic deviation.

Consider any partition of $\boldsymbol{C}$ into, say, K subgroups, indexed by $k = 1,\ldots, K$, $\boldsymbol{C} = (\boldsymbol{c}^1,\ldots, \boldsymbol{c}^K)$. The formula for the GE index when written in decomposable form is the following:

$$I_a(\boldsymbol{C}) = \Sigma_k w_a^k I_a(\boldsymbol{c}^k) + I_a(\boldsymbol{\mu}^1,..., \boldsymbol{\mu}^K), \quad (7)$$

where $w_a^k = [(v^k)^a (p^k)^{1-a}]$; v^k is the share of total citations held by articles in subgroup k; p^k is sub-group k's population share, and $I_a(\boldsymbol{\mu}^1,..., \boldsymbol{\mu}^K)$ is the between-group inequality calculated as if each article in subgroup k received that sub-group's mean citation $\boldsymbol{\mu}^k$. In particular, for the partition of distribution $\boldsymbol{C}$ into Π quantiles, $\boldsymbol{C} = (c^1,\ldots, c^\pi, \ldots, c^\Pi)$, we have:

$$I_a(\boldsymbol{C}) = \Sigma_\pi w_a^\pi I_a(\boldsymbol{c}^\pi) + I_a(\boldsymbol{\mu}^1,..., \boldsymbol{\mu}^\Pi). \quad (8)$$

In order to select some member of the GE family of inequality indicators, we may take into account the following three considerations. Firstly, the weights in the within-group term in expression (7), w_a^k, add up to one only for $a = 0$ and $a = 1$. In any other case, the within-group term will not be a weighted average of the sub-group values $I_a(\boldsymbol{c}^k)$. More importantly, it can be shown that $1 - \Sigma_k w_a^k$ is proportional to the between-group term in (7). This leads to serious difficulties of interpretation of the decomposition in question (see Shorrocks, 1980). Secondly, the behavior of the members of the family when $a \geq 2$ are rather extreme: they show increasingly little concern for transfers except among the very highly cited articles (see also Shorrocks, 1980). In highly skewed distributions this can be problematic. For example, Albarrán *et al.* (2012) show that the elimination of the most highly cited article in each of the 22 fields, that is to say, 22 articles among a dataset of 4.4 million, reduces citation inequality by more than 5%. Thirdly, in the case $a = 2$, for example, the weights become $w_2^k = (\mu^k/\mu)\, v^k$. In particular, for the partition into Π quantiles in expression (8), we have $w_2^\pi = (\mu^\pi/\mu)\, v^\pi$. Thus, for high values of π, w_2^π becomes very high indeed. As we will presently see, the last two facts imply that most of the *IDCP* term is accounted for by the last few quantiles.

These considerations advise choosing either $a = 0$ or $a = 1$. In the first case, for every π, $w_0^\pi = p^\pi = N/\Pi$, that is, w_0^π is quantile's π demographic share. Instead, $w_1^\pi = v^\pi$, the share of citations in quantile π relative to total citations. In economics, the demographic weighting by p^π when $a = 0$ is usually preferred on normative grounds. In our context, the choice $a = 1$ seems more appropriate, in which case the higher the quantile π, the greater the weight v^π assigned to $I_1(\boldsymbol{\mu}_1^\pi,\ldots, \boldsymbol{\mu}_F^\pi)$ in the *IDCP* term in Eq. 5. The problem with this choice (as well as in the case $a = 0$), is that there is a considerable percentage of articles in all fields that receive zero citations, and the index I_1 (as well as I_0) in (6) is only defined for positive numbers. Therefore, we experimented with the following options: assigning to articles without citations the values $\varepsilon_1 = 0.1$, and $\varepsilon_2 = 0.01$ whenever $a = 0, 1$, or adopting the convention $0 \log(0) = 0$ for these articles in the case $a = 1$. Since we must decide at the same time on the value of Π, we have estimated Eq. 5 for the following choices: (i) $a = 2$; (ii) $a = 0$ and $\varepsilon_1 = 0.1$; (iii) $a = 0$ and $\varepsilon_2 = 0.01$; (iv) $a = 1$ and $\varepsilon_1 = 0.1$; (v) $a = 1$ and $\varepsilon_2 = 0.01$, and (vi) $a = 0$ and $0 \log(0) = 0$, on the one hand, and $\Pi = 10, 50, 100$, and 1,000 on the other hand. The results for 22 field case are in Table 1A.

Table 1A around here

The following three points should be noted. Firstly, when a = 2 the quantile choice affects the relative importance of the three terms in decomposition (5). As Π grows, the vectors $\boldsymbol{c_f^{\pi}}$ for all f become smaller and smaller. Consequently, the within-group citation inequality term W loses importance in favor of the S term. Unfortunately, the *IDCP* term is also pretty sensitive to the quantile choice. Secondly, when a = 0 a similar pattern is observed, with two differences: the term W is very small indeed, and the sensitivity of the *IDCP* term to Π is smaller than when a = 2. However, being very sensitive to transfers at the lower tail of citation distributions, the importance of the *IDCP* term according to I_0 is generally higher than for the other two choices of parameter a, and dramatically increases when we go from ε_1 to ε_2. Thirdly, when a = 1 the *IDCP* term remains essentially constant for all choices of Π and ε, ranging from 13.22% to 13.95%. Moreover, the order of magnitude of the *IDCP* term is similar to the case a = 2. Thus, we decide to stick to the following choices for both 22 field and 219 sub-field cases: a = 1, 0 log(0) = 0, and Π = 1,000.

4. 3. Two Strategies for the 219 Sub-field Case

Unlike the case of 22 fields, one difficulty rising from the case of 219 sub-fields is that one article can be assigned into more than one sub-fields. To overcome this problem, two different strategies can be applied, namely the fractional and the multiplicative strategy.

4. 3. 1. The Fractional Strategy

Suppose we have an initial citation distribution $Q = \{c_l\}$ consisting of N distinct articles, *indexed by* $l = 1,\ldots, N$, where c_l is the number of citations received by article l. The total number of citations is denoted by $\gamma = \Sigma_l\, c_l$. Assume that here are S sub-fields, indexed by $s = 1,\ldots, S$. As indicated in the Introduction, the problem is that about 42% of all articles in our dataset are assigned to two or more sub-fields. For later reference, let N_s be the number of distinct articles in sub-field s under the multiplicative approach, indexed by $i = 1,\ldots, N_s$.

For any l, let X_l be the non-empty set of sub-fields to which article l is assigned, and denote by x_l the cardinal of this set, that is, $x_l = |X_l|$. Since, at most, an article is assigned to six sub-fields, we have that $x_l \in [1, 6]$. In the fractional strategy, sub-field s's citation distribution can be described by $c_s = \{w_{si}, c_{si}\}$, where $c_{si} = c_l$ for some article l in the initial distribution Q, $w_{si} = (1/x_l)$ for all $s \in X_l$, and $i = 1,\ldots, N_s$. Therefore, $\sum_{s\in X_l} w_{si} = 1$. The fractional number of articles in sub-field s is $n_s = \Sigma_i\, w_{si}$, the citations received by each fractional article are $w_{si}\, c_{si}$, and the fractional number of citations in sub-field s is $\Sigma_i\, w_{si}\, c_{si}$. It should be noted that $\sum_s n_s = \sum_s \sum_i w_{si} = \sum_l \sum_{s\in X_l} w_{si} = N$ and $\sum_s \sum_i w_{si} c_{si} = \gamma$, that is, in the fractional strategy the total number of articles and citations in the original dataset, and hence the mean citation, are preserved.

Any distinct article i in sub-field s with $c_{si} = c_l$ for some l in the initial distribution $\boldsymbol{Q}$, is assumed to have a scientific influence q_{si} that, for simplicity, is taken to be a single-dimensional variable. We assume that the citations received c_{si} are a function of two variables: the sub-field s to which the article belongs, and the scientific influence of the article in question, q_{si}. Thus, for every s we write:

$$c_{si} = \phi(s, q_{si}),\ i = 1,\ldots, N_s. \tag{9}$$

Let $\boldsymbol{q_s} = \{w_{si}, q_{si}\}$ with $q_{s1} \leq q_{s2} \leq\ldots\leq q_{sN_s}$ be the ordered distribution of scientific influence in every sub-field in the fractional case. Each distribution $\boldsymbol{q_s}$ is assumed to be a characteristic of sub-field s. No restriction is *a priori* imposed on distributions $\boldsymbol{q_s}$, $s = 1,\ldots, S$. Consequently, for any two articles i and j in two different fields s and t the values $w_{si}\, q_{si}$ and $w_{tj}\, q_{tj}$ cannot be directly compared. To overcome this difficulty, we adopt the following key assumption.

Assumption 1b (A1b). *Articles at the same quantile* π *of any sub-field scientific influence distribution have the same degree of scientific influence in their respective field.*

Typically, scientific influence is an unobservable variable. However, although the form of ϕ in Eq. 1 is unknown, we adopt the following assumption about it:

Assumption 2b (A2b). *The function* ϕ *in expression (9) is assumed to be monotonic in scientific influence, that is, for every pair of articles i and j in sub-field s, if* $q_{si} \leq q_{sj}$, *then* $c_{si} \leq c_{sj}$.

Under A2b, the degree of scientific influence uniquely determines the location of an article in its sub-field citation distribution. Consequently, for every s, the partition of distribution $\boldsymbol{q_s}$ into Π quantiles $\boldsymbol{q_s}^{\boldsymbol{\pi}}$ of size n_s/Π, induces a corresponding partition of the citation distribution $\boldsymbol{c_s}$ into Π quantiles $\boldsymbol{c_s}^{\boldsymbol{\pi}}$ with the number of citations received by the n_s/Π articles in the π-th quantile $\boldsymbol{q_s}^{\boldsymbol{\pi}}$. Note that $\boldsymbol{c_s}^{\boldsymbol{\pi}} = \{w_{sk}^{\pi}, c_{sk}^{\pi}\}$, with $c_{sk}^{\pi} = c_{si} = c_l$, and $w_{sk}^{\pi} = 1/x_l$ for some $k = 1,\ldots, N_s$ and some l in $\boldsymbol{Q}$. Assume for a moment that we disregard the citation inequality within every vector $\boldsymbol{c_s}^{\boldsymbol{\pi}}$ by assigning to every article in that vector the (fractional) mean citation of the vector itself, μ_s^{π}, defined by $\mu_s^{\pi} = (\sum_{i\in\pi} w_{si}c_{si})/\sum_{i\in\pi} w_{si}$. Since the quantiles of citation impact correspond –as we have already seen– to quantiles of the underlying scientific influence distribution, holding constant the degree of scientific influence at any π as in A1b is equivalent to holding constant the degree of citation impact at that quantile. Thus, for any π, the difference between μ_s^{π} and μ_t^{π} for articles with the same degree of scientific influence is entirely attributable to differences in citation practices between the two sub-fields. For any s and π, let $\boldsymbol{\mu_s}^{\boldsymbol{\pi}} = \{w_{sk}^{\pi}, \mu_s^{\pi}\}$ be the (n_s/Π)-vector where every c_{sk}^{π} in $\boldsymbol{c_s}^{\boldsymbol{\pi}} = \{w_{sk}^{\pi}, c_{sk}^{\pi}\}$ has been replaced by the mean citation μ_s^{π}. As before, the citation inequality of distribution $(\boldsymbol{\mu_1}^{\boldsymbol{\pi}},\ldots, \boldsymbol{\mu_s}^{\boldsymbol{\pi}},\ldots, \boldsymbol{\mu_S}^{\boldsymbol{\pi}})$ is entirely due to differences in citation practices between the S sub-fields.

To implement our measurement framework, it is convenient to work with additively decomposable citation inequality indices. For reasons explained in Crespo *et al.* (2013a), we choose a member of the so-called Generalized Entropy family of inequality indices, which are the only measures of relative inequality that satisfy the usual properties required from any inequality index and, in addition, are decomposable by population subgroup. This is the first Theil index, denoted by I_1, and defined by:

$$I_1(Q) = (\frac{1}{N}) \sum_l \left(\frac{c_l}{\mu}\right) \log\left(\frac{c_l}{\mu}\right) \quad (10)$$

where μ is the mean of distribution $\boldsymbol{Q}$. Let $\boldsymbol{c}$ be the union of all sub-field distributions $\boldsymbol{c_s}$, that is, let $\boldsymbol{c} = \cup_s \boldsymbol{c_s}$. As we have seen already, the number of articles and the mean citation of distributions Q and $\boldsymbol{c}$ coincide. Clearly, citation inequality is also the same, that is, $I_1(\boldsymbol{c}) = I_1(\boldsymbol{Q})$. Therefore, in the sequel we will work with distribution $\boldsymbol{c}$.

For each π, let $\boldsymbol{c}^{\boldsymbol{\pi}} = (\boldsymbol{c_1}^{\boldsymbol{\pi}},\dots, \boldsymbol{c_s}^{\boldsymbol{\pi}},\dots, \boldsymbol{c_S}^{\boldsymbol{\pi}})$. Note that the vector $\boldsymbol{c}^{\boldsymbol{\pi}}$ has dimension $\Sigma_s (n_s/\Pi) = N/\Pi$, and that the set $\boldsymbol{c}^{\boldsymbol{\pi}}$, $\pi = 1,\dots, \Pi$, form a partition of distribution $\boldsymbol{c}$. For any π, let $\boldsymbol{\mu}^{\boldsymbol{\pi}}$ be the (N/Π)-vector where every element in $\boldsymbol{c}^{\boldsymbol{\pi}}$ has been replaced by the mean citation $\mu^{\pi} = \Sigma_s [(n_s/N]\mu_s^{\pi}$. As in Crespo *et al.* (2013a), applying the decomposability property of citation inequality index I_1 first to the partition $\boldsymbol{c} = (\boldsymbol{c}^{\boldsymbol{1}},\dots, \boldsymbol{c}^{\boldsymbol{\pi}},\dots, \boldsymbol{c}^{\boldsymbol{\Pi}})$, and then to the partition $\boldsymbol{c}^{\boldsymbol{\pi}} = (\boldsymbol{c_1}^{\boldsymbol{\pi}},\dots, \boldsymbol{c_s}^{\boldsymbol{\pi}},\dots \boldsymbol{c_S}^{\boldsymbol{\pi}})$ for each π, the overall citation inequality $I_1(\boldsymbol{c})$ can be seen to be decomposable into the following three terms:

$$I_1(\boldsymbol{c}) = W + S + IDCP, \quad (11)$$

with:

$$W = \Sigma_{\pi} \Sigma_s v^{\pi,s} I_1(\boldsymbol{c_s}^{\boldsymbol{\pi}})$$

$$S = I_1(\boldsymbol{\mu}^{\boldsymbol{1}},\dots, \boldsymbol{\mu}^{\boldsymbol{\Pi}})$$

$$IDCP = \Sigma_{\pi} v^{\pi} I_1(\boldsymbol{\mu_1}^{\boldsymbol{\pi}},\dots, \boldsymbol{\mu_S}^{\boldsymbol{\pi}}) = \Sigma_{\pi} v^{\pi} I(\pi),$$

where $v^{\pi,s}$ is the share of total citations in quantile π of sub-field s, and $v^{\pi} = \Sigma_s v^{\pi,s}$ is the share of total citations in vector $\boldsymbol{c}^{\boldsymbol{\pi}}$. The term W is a within-group term that captures the weighted citation inequality within each quantile in every sub-field. For large Π, W is expected to be small. The term S is the citation inequality of the distribution $(\boldsymbol{\mu}^{\boldsymbol{1}},\dots, \boldsymbol{\mu}^{\boldsymbol{\Pi}})$, and therefore it is a measure of citation inequality at different degrees of citation impact in the all-sciences case. Due to the skewness of science, S is expected to be large. Finally, for any π, the expression $I_1(\boldsymbol{\mu_1}^{\boldsymbol{\pi}},\dots, \boldsymbol{\mu_S}^{\boldsymbol{\pi}})$, abbreviated as $I(\pi)$, is the citation inequality attributable to differences in citation practices according to I_1. Thus, the weighted average that constitutes the third term in expression (11), denoted by *IDCP* (*I*nequality due to *D*ifferences in *C*itation *P*ractices), provides a good measure of the citation inequality due to such differences at the sub-field level. The question of interest, of course, is how large is the *IDCP* term in relation to overall citation inequality $I_1(\boldsymbol{c})$.

4. 3. 2. The Multiplicative Strategy

In the multiplicative approach each article is wholly counted as many times as necessary in the several sub-fields to which it is assigned. In this way, the space of articles is expanded as much as necessary beyond the initial size in what we call the *sub-field extended count*, say distribution $\boldsymbol{C}$. In this approach, sub-field s's citation distribution can be described by $C_s = \{c_{si}\}$ with $i = 1,\dots, N_s$, where c_{si} is the number of citations of article i in sub-field s, and $c_{si} = c_l$ for some article l in the initial distribution $\boldsymbol{Q}$. Of course, $\boldsymbol{C} = \cup_{\boldsymbol{s}} \boldsymbol{C_s}$, and the total number of articles in the sub-field extended count is $M = \Sigma_s N_s > N$.

In what follows, let us order sub-field citation distributions, so that for any s we have $C_s = (c_{s1},\dots, c_{si},\dots, c_{sN_s})$ with $c_{s1} \leq c_{s2} \leq \dots \leq c_{sN_s}$. Consider the partition of distribution C_s into Π quantiles, $C_s = (C_s^1,\dots, C_s^{\pi},\dots, C_s^{\Pi})$, where each vector $C_s^{\pi} = \{c_{sj}\}$ with $j = 1,\dots, N_s/\Pi$. For each π, define the citation distribution $C^{\pi} = (C_1^{\pi},\dots, C_s^{\pi},\dots, C_S^{\pi})$. Clearly, the number of articles in C^{π} is $\Sigma_s N_s/\Pi = M/\Pi$, and the set of vectors $(C^1,\dots,C^{\pi},\dots, C^{\Pi})$ form a partition of distribution $\boldsymbol{C}$. For any s and π, let $\boldsymbol{m_s^{\pi}}$ be the (N_s/Π)-vector where every c_{sj}^{π} in $\boldsymbol{C_s^{\pi}} = \{c_{sj}^{\pi}\}$ has been replaced by the mean citation $m_s^{\pi} = (\Sigma_j c_{sj}^{\pi})/(N_s/\Pi)$. Similarly, for any π, let $\boldsymbol{m^{\pi}}$ be the (N/Π)-vector where every element in $\boldsymbol{C^{\pi}}$ has been replaced by the mean citation $m^{\pi} = \Sigma_{\boldsymbol{s}}(n_{\boldsymbol{s}}/N)m_s^{\pi}$. Applying the decomposability property of citation inequality index I_1 first to the partition $C = (C^1,\dots,C^{\pi},\dots, C^{\Pi})$, and then to the partition $C^{\pi} = (C_1^{\pi},\dots, C_s^{\pi},\dots, C_S^{\pi})$ for each π, the overall citation inequality $I_1(\boldsymbol{C})$ can be seen to be decomposable into the following three terms analogous to what we had in expression (11):

$$I_1(\boldsymbol{C}) = W' + S' + IDCP', \tag{12}$$

with:

$$W' = \Sigma_{\pi} \Sigma_s V^{\pi,s} I_1(\boldsymbol{C_s^{\pi}})$$

$$S' = I_1(\boldsymbol{m^1},\dots, \boldsymbol{m^{\Pi}})$$

$$IDCP' = \Sigma_{\pi} V^{\pi} I_1(\boldsymbol{m_1^{\pi}},\dots, \boldsymbol{m_S^{\pi}}),$$

where $V^{\pi,s}$ is the share of total citations in quantile π of sub-field s, and $V^{\pi} = \Sigma_s V^{\pi,s}$ is the share of total citations in vector $\boldsymbol{C^{\pi}}$. As before, the term W' is a within-group citation inequality term, S' captures the skewness of science, and $IDCP'$ is the citation *I*nequality that can be attributed to *D*ifferences in *C*itation *P*ractices in the multiplicative case.

4. 3. 3. Empirical Results of the 219 Sub-field Case

The original dataset consists of 4.4 million articles published in 1998-2003 and 35 million citations they receive after a common five-year citation window for every year.[1] The extended count is 7,027,037, or 57.4% larger than the total number of articles in the fractional approach. Table A2 in the Appendix presents the number of articles and mean citation rates in the fractional case. For convenience, sub-fields are classified in terms of 19 fields, and four large groups: Life Sciences, Physical Sciences, Other Natural Sciences, and Social Sciences, which represent, respectively, 40.1%, 30.2%, 25.8%, and 3.9% of all articles (the same information for the multiplicative case is available on request).

Table 1B, which includes the decompositions of $I_1(\boldsymbol{c})$ and $I_1(\boldsymbol{C})$ presented in expressions (3) and (4) for the value $\Pi = 1{,}000$, deserves the following three comments.[2] Firstly, as in Crespo *et al.*

(2013a), the terms *W* and *W'* are small, while the terms *S* and *S'* are large. Secondly, the importance of the effect on overall citation inequality of differences in citation practices is larger when working with 219 sub-fields than with 22 broad fields. In particular, the *IDCP* term that represents in Crespo *et al.* (2013a) about 14% of overall citation inequality increases four percentage points, up to 17.95%, in the fractional case. Thirdly, interestingly enough the *IDCP'* term in the multiplicative case represents 18.1% of overall citation inequality, a figure remarkably close to the corresponding one in the fractional case.

Table 1B around here

5. COMPARABILITY AND NORMALIZATION PROCEDURES

This Section analyzes two empirical problems: (i) how to compare the citations received by two articles in any pair of the 22 fields (219 sub-fields) in our dataset by using exchange rates that are approximately constant over a large quantile interval, and (ii) how much the effect of differences in citation practices is reduced when these exchange rates, or the field (sub-field) mean citations are used as normalization factors.

5. 1. The Case of 22 Fields

5. 1. 1. The Comparison of Citation Counts across Different Fields

Mean citations of comparable articles belonging to the same quantile can be used to express the citations in any field in terms of the citations in a reference situation. For example, if we let μ^{π} be the mean citation of all articles in quantile π, then the *exchange rates at quantile* π, $e_f(\pi)$, defined by

$$e_f(\pi) = \mu_f^{\pi} / \mu^{\pi}, \tag{13}$$

can be seen to answer the following question: how many citations for an article at the degree π of scientific influence in field *f* are equivalent on average to one citation in the all-fields case? In the metaphor according to which a field's citation distribution is like an income distribution in a certain currency, the exchange rates $e_f(\pi)$ permit to express all citations in the same reference currency for that π: since c_{fi} is the number of citations received by article *i* in quantile π of field *f*, the ratio $c_{fi}^*(\pi) = c_{fi} / e_f(\pi)$ is the equivalent number of citations in the reference currency at that quantile. Naturally, if for many fields $e_f(\pi)$ were to drastically vary with π, then we might not be able to claim that differences in citation practices have a common element that can be precisely estimated. However, we next establish that exchange rates are sufficiently constant over a wide range of quantiles.

It is very instructive to have a graphical representation of how the effect of differences in citation practices, measured by $I(\pi)$, changes with π when $\Pi = 1{,}000$ (since $I(\pi)$ is very high for $\pi < 600$, for clarity these quantiles are omitted from Figure 1A). It is observed that $I(\pi)$ is particularly high until $\pi \approx 700$, as well as for a few quantiles at the very upper tail of citation distributions. However,

$I(\pi)$ is strikingly similar for a wide range of intermediate values.[6] In this situation, it is reasonable to define an average-based *exchange rate* (*ER* hereafter) over some interval $[\pi_m, \pi^M]$ in that range as

$$e_f = [1/(\pi^M - \pi_m)] [\Sigma_\pi e_f(\pi)]. \tag{14}$$

An advantage of this definition is that we can easily compute the associated StDev, denoted by σ_f. The fact that, for each f, the $e_f(\pi)$ defined in (13) are very similar for all π in the interval $[\pi_m, \pi^M]$ would manifest itself in a small σ_f, and hence in a small coefficient of variation $CV_f = \sigma_f / e_f$.

Figure 1A around here

We find that the choice $[\pi_m, \pi^M] = [706, 998]$ –where $I(\pi)$ for most π is equal to or smaller than $I(\pi_m) = 0.1081$ and $I(\pi^M) = 0.1084$– is a good one. The *ER*s e_f, as well as the σ_f, and CV_f are in columns 1 to 3 in Table 2A. For convenience, *ER*s are multiplied by 10. Thus, for example, the first row indicates that 15.8 citations with a StDev of 0.9 for an article in Biology and Biochemistry between, approximately, the 71st and the 99th percentile of its citation distribution, are equivalent to 10 citations for an article in that interval in the all-sciences case. We find it useful to divide fields into three groups according to the CV_f. Group I (colored in green in Table 1A), consisting of 10 fields, has a CV_f smaller than or equal to 0.05. This means that the StDev of the exchange rate is less than or equal to five percent of the exchange rate itself. Hence, we consider *ER*s in this group as highly reliable. Group II (black), consisting of 10 fields, has a CV_f between 0.05 and 0.10. We consider *ER*s in this group as fairly reliable. Group III (red), consists of two fields: *Computer Science*, with a CV_f greater than 0.10, which is known from previous work to behave as an outlier (Herranz and Ruiz-Castillo, 2012), and the *Multidisciplinary* field with a CV_f greater than 0.15, a hybrid field that does not behave well either in RC. The results for these two fields should be considered unreliable.

Table 2A around here

As is observed in column 4 in Table 2A, on average the interval [706, 998] includes 72.1% of all citations (with a StDev of 3.9). Expanding the interval in either direction would bring a larger percentage of citations. It turns out that the *ER*s do not change much. However, they exhibit greater variability. For example, moving the upper bound π^M to quantile 1,000 would increase the percentage of citations to 76.3% (StDev = 5). However, the CV_f would increase in all but three fields, and the number of fields in Group I would decrease from 10 in the reference case down to 8. In the other direction, moving the lower bound π_m to quantiles 700, or 694, for example, would slightly increase the percentage of citations to 72.7%, (StDev = 3.8) and 73.3% (StDev = 3.8). However, relative to the initial choice, in these two instances the CV_f would increase in 13 out of 22 fields, and the number of fields in Groups I would decrease from 10 to 9. On the other hand, after normalization by the *ER*s

[6] It is important to emphasize that this is consistent with the stylized facts characterizing citation distributions discussed in Section II and documented in Albarrán and Ruiz-Castillo (2011), and Albarrán *et al.* (2012): although the percentages of articles belonging to three broad classes are very similar across fields, citation distributions are rather different in a long lower tail and at the very top of the upper tail.

corresponding to the three alternatives [706, 1000], [700, 998], and [694, 998], the *IDCP* term represents essentially the same percentage of the overall citation inequality in the normalized distributions (see below). Therefore, we retain the interval [706, 998] in the sequel.

5. 1. 2. Normalization Results

Overall citation inequality due to differences in scientific influence –captured by the W and S terms in Eq. 6– is not worrisome. Instead, we would like to eliminate as much as possible the citation inequality attributable to differences in citation practices. Thus, the impact of any normalization procedure can be evaluated by the reduction in the term *IDCP* before and after normalization. Figure 2A focuses on the product $v^{\pi} I(\pi)$ as a function of π. Of course, the term *IDCP* is equal to the integral of this expression (for clarity, quantiles $\pi < 600$, and $\pi > 994$, are omitted from Figure 2A). Note the strong effect of the weights v^{π} as π increases. As a matter of fact, the percentage of *IDCP* reached at $\pi = 400$, 700, 900, and 990 are 15.2%, 35.9%, 61.9%, and 88.9%, respectively.[7]

Relative to the blue curve, the red curve illustrates the correction achieved by normalization: the size of the *IDCP* term is very much reduced. The numerical results before and after this normalization are in Panels A and B in Table 3A. Note that, as before, the term W is small, while the term S is large. Both terms remain essentially constant after normalization. However, in absolute terms the *IDPC* term is reduced from 0.1221 to 0.0167, a 86.3% difference. Of course, total citation inequality after normalization is also reduced. On balance, the *IDPC* term after normalization only represents 2.09% of total citation inequality –a dramatic reduction from the 13.95% with the raw data.

Table 3A and Figure 2A around here

However, it should be recognized that in the last two quantiles and, above all, in the [1, 705] interval normalization results quickly deteriorate. It would appear that a convenient alternative consists of normalizing the lower tail of the original distributions by some appropriate *ER*s within the [1, 705] interval. The problem is that citation inequality due to different citation practices in that interval is both high and extremely variable for different quantiles. It turns out that the *ER*s computed according to equation (14) for the entire [1, 705] interval lead to a worsening of the situation. However, when we restrict ourselves to the interval [356, 705] we are able to improve matters somewhat. The new *ER*s, together with their high σ_f, and CV_f, are in Table B1 in the Appendix. The second set of *ER*s is rather different: only in seven cases do they stay within one StDev of the first set in Table 2A. On the other hand, CV_fs increase so much that seven fields are now in Group IV when we only had one field in that group before. Be that as it may, the end result is that after normalization by the two sets of *ER*s the *IPC* only goes down to 1.86% of total citation inequality (see Panel C in Table 3A) versus 2.09% with a single set of *ER*s. Most of this figure, or 1.36%, is still accounted for by what happens in the interval

[7] Being very sensitive to transfers at the upper tail of citation distributions, the percentage of *IDCP* reached at $\pi = 400$, 700, 900, and 990 according to I_2 are very much lower than according to I_1: 0.7%, 4.7%, 17.1%, and 50%, respectively. Thus, half of the *IDCP* is accounted for by the last ten quantiles.

[1, 705]. We must conclude that the improvement over the alternative with a single set of *ER*s is, at most, very slight.

As indicated in the Introduction and discussed in Section II, the difficulties of combining heterogeneous citation distributions into broader aggregates have been traditionally confronted using mean citations as normalization factors. In our dataset, the *IDCP* term after the traditional normalization procedure only represents 2.05% of total citation inequality (see Panel D in Table 3A). The two solutions are so near that we refrain to illustrate the latter in Figure 2A because it will be indistinguishable with the red curve after normalization by our *ER*s. This confirms the results in RC, where it is concluded that, despite not being strictly correct, this procedure is a very good approximation of the two-parameter transformation able to make citation counts independent of the scientific field.

The question is, how can this similarity of results be accounted for? The explanation is as follows. As documented in Albarrán *et al.* (2011), field mean citations μ_f are reached, on average, at the 69.7 percentile with a StDev of 2.6, that is, at the lower bound of the [706, 998] interval. Thus, the *ER*s based on mean citations, $e_f(\mu_f) = \mu_f/\mu$ (reproduced in column 5 in Table 2A), are approximately equal our own *ER*s (in column 1 in Table 2A). In other words, let μ'_f and μ' be the mean citations in each field and the population as a whole restricted to the [706, 998] interval, and consider the average-based *ER*s based on these restricted means: $e_f(\mu'_f) = \mu'_f/\mu'$ (see column 6 in Table 2A). Since field citation distributions differ approximately by a set of scale factors only in the interval [706, 998], these scale factors should be well captured by any average-based measure of what takes place in that interval –such as our own e_f, or the new $e_f(\mu'_f)$. However, the latter *ER*s are essentially equal to the old ones, that is, for each f, $e_f(\mu'_f) \approx e_f(\mu_f) \approx e_f$.

Finally, we have estimated the reduction in the *IDCP* term when, following Glänzel (2011), the normalization factors are made equal to the difference $(s_2 - s_1)$ for each field, where s_1 and s_2 are the first two scores in the Characteristic Scores and Scales approach discussed in Section II. The results are in Panel E of Table 3A. Interestingly, for the last two quantiles the reduction is larger than in all previous cases. However, the entire *IDCP* term after this third normalization becomes 3% –rather than 2%– of overall citation inequality.

5. 2. The Case of 219 Sub-fields

5. 2. 1. The Fractional Strategy

This Section analyzes two empirical problems in the fractional case: (i) how to compare the citations received by two articles in any pair of the 219 sub-fields in our dataset by using *ER*s that are approximately constant over a large quantile interval, (ii) how much the *IDCP* term is reduced when these *ER*s, or the field mean citations are used as normalization factors. The robustness of these results in the multiplicative approach is studied in Section IV.

5. 2. 1. 1. The Comparison of Citation Counts across Different Sub-fields

For any s, what we call the *exchange rates at quantile* π, $e_s(\pi)$, defined by

$$e_s(\pi) = \mu_s^{\pi} / \mu^{\pi}, \qquad (15)$$

allows us to answer the following question: how many citations for an article at the degree π of scientific influence in sub-field s are equivalent on average to one citation in the all-fields case? In the metaphor according to which a sub-field's citation distribution is like an income distribution in a certain currency, the exchange rates $e_s(\pi)$ permit to express all citations in the same reference currency for that π.

Naturally, if for many fields $e_s(\pi)$ were to drastically vary with π, then we might not be able to claim that differences in citation practices have a common element that can be precisely estimated. However, it has been established that the shapes of sub-field citation distributions are highly skewed and, what is more important for our purposes, very similar indeed.[3] As we will presently see, the similarity between sub-field citation distributions imply that exchange rates are sufficiently constant over a wide range of quantiles.

Figure 1B represents how the effect of differences in citation practices, measured by $I(\pi)$, changes with π when $\Pi = 1{,}000$ (since $I(\pi)$ is very high for $\pi < 260$, for clarity these quantiles are omitted from Figure 1B). It is observed that $I(\pi)$ is particularly high until $\pi \approx 600$, as well as for a few quantiles at the very upper tail of citation distributions. However, as in Crespo *et al.* (2013a) $I(\pi)$ is rather similar for a wide range of intermediate values, indicating that, over that interval, sub-field citation distributions essentially differ by a scale factor. In this situation, for each s it is reasonable to define an average-based *exchange rate* (ER) over some interval $[\pi_m, \pi^M]$ in that range as

$$ER_s = [1/(\pi^M - \pi_m)] [\Sigma_{\pi} e_s(\pi)], \qquad (16)$$

where, for each π,

$$e_s(\pi) = \mu_s^{\pi} / \mu^{\pi}.$$

Figure 1B around here

We find that the choice $[\pi_m, \pi^M] = [661, 978]$ –where $I(\pi)$ for most π is equal to $I(\pi_m) = 0.1356$ and $I(\pi^M) = 0.1392$– is a good one. The ERs, as well as the StDev, and the coefficient of variation (CV hereafter) are in columns 1 to 3 in Table 2B. For convenience, ERs are multiplied by 10. Thus, for example, the first row indicates that 10.3 citations with a StDev of 0.3 for an article in Biology between, approximately, the 66st and the 98th percentile of its citation distribution, are equivalent to 10 citations for an article in that interval in the all-sciences case. We find it useful to divide fields into four groups according to the CV. Group I (colored in dark green in Table 2B), consisting of 69 sub-fields, has a CV smaller than or equal to 0.05. This means that the StDev of the exchange rate is less than or equal to five percent of the exchange rate itself. Hence, we consider ERs in this group as highly reliable. Group II (pale green), consisting of 118 sub-fields, has a CV between 0.05 and 0.10. We consider ERs in this group as fairly reliable. Group III (orange), consists of 22 sub-fields, has a CV between 0.10 and 0.15. This group includes some important sub-fields, such as *Physics, Particles and Fields; Information and Library Science*, and *Political Science* (sub-fields 97, 210, and 189), as well as seven out of eight sub-fields within the broad field *Computer Science* (the exception is *Mathematical and Computational Biology*) that is known to behave as an outlier (Herranz and Ruiz-Castillo, 2012, and Crespo *et al.*, 2013a). Some would find *ER*s in this group as minimally reliable, while others will find them quite unreliable. Finally, Group IV (red), consisting of nine sub-fields, has a *CV* greater than 0.15. This group includes *Multidisciplinary Sciences* and *Physics, Multidisciplinary*, hybrid sub-fields some of which also behave badly in Radicchi and Castellano (2012a). *ER*s in this group can be considered unreliable.

Table 2B around here

As is observed in column 4 in Table 2B, on average the [661, 978] interval includes 62.2% of all citations (with a StDev of 3.0). Although this is a relatively large percentage, expanding the interval in either direction would bring a larger percentage of citations. It turns out that, when we do this, the *ER*s do not change much. However, they exhibit greater variability (see the details in Crespo *et al.*, 2013b). Therefore, we retain the interval [661, 978] in the sequel.

5. 2. 1. 2. Normalization Results

In the first place, we want to assess the normalization procedure based on *ER*s whereby the citations received by any article i in sub-field s, c_{si}, are converted into normalized citations c_{si}^* as follows: $c_i^* = c_{si}/ER_s$. The numerical results before and after this normalization are in Panels A and B in Table 3B. As in Crespo *et al.* (2013a), the terms W and S remain essentially constant after normalization by the *ER*s. In absolute terms the *IDPC* term is reduced from 0.1552 to 0.0293, a 81.1% difference. Of course, total citation inequality after normalization is also reduced. On balance, the *IDPC* term after normalization only represents 3.85% of total citation inequality –an important reduction from the 17.95% with the raw data.

Table 3B around here

However, it should be recognized that in the last 22 quantiles and, above all, in the [1, 660] interval normalization results quickly deteriorate. Figure 2B, which focuses on the product $v^{\pi} I(\pi)$ as a function of π, illustrates the situation. Of course, the term *IDCP* introduced in expression (11) is equal to the integral of this expression (for clarity, quantiles $\pi < 600$, and $\pi > 994$, are omitted from Figure 2B). Relative to the blue curve, the red curve illustrates the correction achieved by normalization with the 219 *ER*s: the size of the *IDCP* term is very much reduced, particularly in the [661, 978] interval.

Figure 2B around here

Finally, as in Crespo *et al.* (2013a) it is interesting to examine the consequences of the traditional procedure in which sub-field mean citations are taken as normalization factors. The exchange rates based on mean citations, $e_s(\mu_s) = \mu_s/\mu$ (see column 5 in Table 2B) are very close indeed to our own ER_s (see Figure 3B for an illustration). As a matter of fact, they are between one StDev of the ER_s for 50 out of 69 sub-fields in Group I, 102 out of 118 in Group II, 22 out of 23 in Group III, and in all nine cases in Group IV. When sub-field mean citations are used as normalization factors, the *IDCP* term only represents 3.45% of total citation inequality (see Panel C in Table 3B). The two solutions are so near that we refrain to illustrate the latter in Figure 2B because it will be indistinguishable from the red curve after normalization by our *ER*s.[4]

Figure 3B around here

The similarity between the results of the two normalization procedures lies in the fact that, as we have seen in Figure 1B, sub-field citation distributions appear to differ by a set of scale factors only in the [660, 978] interval. These scale factors are well captured by any average-based measure of what takes place in that interval –such as our *ER*s. However, as indicated in note 3, sub-field mean citations in the fractional approach, μ_s, are reached, on average, at the 68.3 percentile with a StDev of 3.4, that is, in the interior of the [661, 978] interval. This is the reason why the *ER*s based on mean citations also work so well.

5. 2. 2. The Multiplicative Strategy

The information about the evolution of $I(\pi)$ as a function of π (available on request), as well as the aim of facilitating the comparison with the fractional case justifies the same choice as before: $[\pi_m, \pi^M]$ = [661, 978]. The corresponding *ER*s, StDevs, and *CV*s are in columns 1 to 3 in Table 4B. As observed in column 4, on average the percentage of citations covered in this interval is 62.3% (with a StDev equal to 3.0). The *ER*s based on sub-field citation means appear in column 5, while the consequences of the normalization using both sets of ERs are in Table 5B.

Tables 4B and 5B around here

The massive information deserves the following four comments. Firstly, Groups I, II, III, and IV consist now of 77, 113, 19, and 10 sub-fields –figures that slightly improve on those obtained in the fractional case. Secondly, the normalization using our own *ER*s or those based on sub-field mean citations reduces the *IDCP'* term to 3.57% and 3.27%, respectively. Thus, in both cases normalization results slightly improve what was obtained under the fractional approach. Thirdly, it should be emphasized that the success of our empirical strategy in the multiplicative case is again based on the similarity of the shapes of sub-field citation distributions.[5] Fourthly, the results in the fractional and the multiplicative cases are extremely similar: except for two sub-fields, the multiplicative *ER*s are always within one StDev of the fractional ones (see Figure 4B for an illustration). As indicated in Herranz and Ruiz-Castillo (2012), the similarity of the citation characteristics of articles published in journals assigned to one or several sub-fields guarantees that choosing one of the two strategies may not lead to a radically different picture in practical applications.

Figure 4B around here

6. CONCLUSIONS

The lessons that can be drawn from this paper can be summarized as follows.

(1) We have provided a simple method for the measurement of the effect of differences in citation practices across scientific fields. Using a member of a family of additively separable citation inequality indices, this effect is well captured by a between-group term –denoted *IDCP*– in the double partition by field and quantile of the overall citation distribution in the all-fields case. It should be noted that this is a distribution free method, in the sense that it does not require that the scientific influence or the citation distributions satisfy any specific assumptions. We use a large dataset of 4.4 million articles and a five-year citation window. When the classification of articles goes from 22 broad fields to 219 sub-fields, the estimated IDCP term increases. We have estimated that the *IDCP* term represents about 14% of overall citation inequality in the case of 22 fields and about 18% in the case of 219 sub-fields.

(2) The striking similarity of citation distributions allows the effect of idiosyncratic citation practices to be rather well estimated over a wide range of intermediate quantiles where citation distributions seem to differ by a scale factor. Consequently, a set of *ER*s has been estimated in the interval [706, 998] for the case of 22 fields and in the interval [661, 978] for the case of 219 sub-fields. With the ERs, we can translate citation counts of articles in different fields within that interval into the citations in a reference situation, and normalize the raw citation data. Such *ER*s are estimated with a reasonably low StDev for 20 out of 22 fields and 187 out of 219 sub-fields.

It should be stressed that, for uncited and poorly cited articles below the mean, and for articles in the very upper tail of citation distributions, no clear answer to the comparability of citation counts for articles in different fields can be provided. Since the citation process evolves at different velocity in different fields, using variable citation windows to ensure that the process has reached a similar stage in all fields should improve field comparability at the lower tail of citation distributions.

(3) The success of any normalization procedure in eliminating as much as possible the impact of differences in citation practices can be evaluated by the reduction it induces in the *IDCP* term. In our case, it has been established that both the procedure that uses our *ER*s, as well as the traditional method of taking the field citation means as normalization factors reduces the importance of the *IDCP* term relative to overall citation inequality from around 14% to 2% in the case of 22 fields, and from around 18% to 3.8% (3.4% with sub-field mean citations) in the case of 219 sub-fields. The paper provides an empirical explanation of why the two methods are equally successful.

Other normalization proposals –such as the one in RC, or those based on citing side procedures quoted in the Introduction, might be analogously evaluated. In turn, it would be interesting to evaluate the normalization procedure based on the *ER*s in terms of the reduction of the bias in the RC model. Given how near our *ER*s are from those based on the fields' mean citation rates, the conjecture is that our procedure would perform as well as the approximation provided by these means in RC.

(4) Interestingly enough, our results at the lowest aggregate level about the *ER*s and their role as normalization factors in the fractional case are essentially replicated when we adopt the multiplicative approach.

One limitation of this study is that we cannot take into account possible differences in citation practices within sub-fields. For example, large differences between basic and clinical research areas within medical Web of Science subject-categories have been recently revealed in Van Eck *et al.* (2012). Naturally, our methods can be equally applied to future classification systems consisting of more homogeneous sub-fields than the Web of Science constructs available to us in this paper.

(5) Policy makers and other interested parties should be very cautious when comparing citation performance in different scientific fields. More research is still needed. However, together with the important contribution by RC, the results of this paper indicate that the combination of interesting assumptions with the empirical similarity of citation distributions paves the way for meaningful comparisons of citation counts across heterogeneous fields.

REFERENCES

Albarrán, P. and J. Ruiz-Castillo (2011), "References Made and Citations Received By Scientific Articles", *Journal of the American Society for Information Science and Technology*, **62**: 40-49.

Albarrán, P., J. Crespo, I. Ortuño, and J. Ruiz-Castillo (2011), "The Skewness of Science In 219 Sub-fields and a Number of Aggregates", *Scientometrics*, **88**: 385-397.

Albarrán, P., Li, Y., and Ruiz-Castillo, J. (2012), "The Role of Extreme Observations in Citation Analysis", mimeo.

Bornmann, L, and H-D. Daniel (2008), "What Do Citation Counts Measure", *Journal of Documentation*, **64**: 45-80.

Bourguignon, F. (1979), "Decomposable income inequality measures", *Econometrica*, **47**: 901-920.

Braun, T., W. Glänzel, and A. Schubert (1985), *Scientometrics Indicators. A 32 Country Comparison of Publication Productivity and Citation Impact*, World Scientific Publishing Co. Pte. Ltd., Singapore, Philadelphia.

Castellano, C., Li, Y., Radicchi, F., and Ruiz-Castillo, J. (2013), "Quantitative Evaluation of Alternative Field Normalization Procedures", mimeo, Universidad Carlos III of Madrid.

Crespo, J. A., Li, Yunrong, and Ruiz-Castillo, J. (2013a), "The Measurement of the Effect On Citation Inequality of Differences In Citation Practices Across Scientific Fields", *Plos One*, **7**(3): e33833.

Crespo, J. A., Li, Yunrong, Herranz, N., and Ruiz-Castillo, J. (2013b), "The Effect On Citation Inequality Of Differences In Citation Practices At the Web Of Science Subject Category Level", Working Paper, 13-03, Universidad Carlos III (http://hdl.handle.net/10016/16327).

Cowell, F. A. and K. Kuga (1981), "Inequality Measurement: An Axiomatic Approach," *European Economic Review,* **15**: 287-305.

Herranz, N. & Ruiz-Castillo, J. (2012), "Multiplicative and Fractional Strategies When Journals Are Assigned to Several Sub-fields", in press *Journal of the American Society for Information Science and Technology* (DOI:10.1002/asi.22629).

Garfield, E. (1979), *Citation Indexing: Its Theory and Applications in Science, Technology, and Humanities*, New York: Wiley.

Glänzel, W. (2011), "The Application of Characteristic Scores and Scales to the Evaluation and ranking of Scientific Journals", *Journal of Information Science*, **37**: 40-48.

Glänzel, W., Schubert, A., Thijs, B., & Debackere, K. (2011) A priori vs. a posteriori normalization of citation indicators. The case of journal ranking. *Scientometrics*, 87, 415-424.

Leydesdorff, L., and Opthof, T. (2010), "Normalization at the Field level: Fractional Counting of Citations", *Journal of Informetrics*, **4**: 644-646.

Leydesdorff, L., Radicchi, F., Bornmann, L., Castellano, C., and de Nooye, W. (2012), "Field-normalized Impact Factors: A Comparison of Rescaling *versus* Fractionally Counted IFs", in press, *Journal of the American Society for Information Science and Technology.*

Moed H. F. (2010), "Measuring contextual citation impact of scientific journals", *Journal of Informetrics*, **4**: 265–77.

Moed, H. F., W. J. Burger, J. G. Frankfort, and A. F. J. van Raan (1985), "The Use of Bibliometric Data for the Measurement of University Research Performance", *Research Policy*, **14**: 131-149.

Moed, H. F., and A. F. J. van Raan (1988), "Indicators of Research Performance", in A. F. J. van Raan (ed.), *Handbook of Quantitative Studies of Science and Technology*, North Holland: 177-192.

Moed, H. F., R. E. De Bruin, and Th. N. van Leeuwen (1995), "New Bibliometrics Tools for the Assessment of national Research Performance: Database Description, Overview of Indicators, and First Applications", *Scientometrics*, **33**: 381-422.

Murugesan, P., and Moravcsik, M.J. (1978), "Variation of the Nature of Citation Measures with Journal and Scientific Specialties", *Journal of the American Society for Information Science and Technology*, **105**: 17268-17272.

Pinski, G. and Narin, F. (1976), "Citation Influencefor Journal Aggregates of Scientific Publications: Theory, with Applications to the Literature of Physics", *Information Processing and Management*, **12**: 297-312.

Radicchi, F., Fortunato, S., and Castellano, C. (2008), "Universality of Citation Distributions: Toward An Objective Measure of Scientific Impact", *PNAS*, **105**: 17268-17272.

Radicchi, F., and Castellano, C. (2012a), "A Reverse Engineering Approach to the Suppression of Citation Biases Reveals Universal Properties of Citation Distributions", *Plos One*, **7**, e33833, 1-7.

Radicchi, F., and Castellano, C. (2012b), "Testing the fairness of citation indicators for comparisons across scientific domains: The case of fractional citation counts", *Journal of Informetrics*, **6**: 121-130.

Ravallion, M., and Wagstaff, A. (2011), "On measuring scholarly influence by citations", *Scientometrics*, **88**: 321-337.

Roemer, J. (1998), *Equality of Opportunities*, Harvard University Press.

Ruiz-Castillo, J. (2003), "The Measurement of Inequality of Opportunities", in J. Bishop y Y. Amiel (eds.), *Research in Economic Inequality,* **9**: 1-34.

Shorrocks, A. F. (1980), "The Class of Additively Decomposable Inequality Measures," *Econometrica,* **48**: 613-625.

Shorrocks, A. F. (1984), "Inequality Decomposition by Population Subgroups," *Econometrica,* **52**: 1369-1388.

Schubert, A., and T. Braun (1986), "Relative Indicators and Relational Charts for Comparative Assessment of Publication Output and Citation Impact", *Scientometrics*, **9**: 281-291.

Schubert, A., and T. Braun (1996), "Cross-field Normalization of Scientometric Indicators", *Scientometrics*, **36**: 311-324.

Schubert, A., W. Glänzel, and T. Braun (1983), "Relative Citation Rate: A New Indicator for Measuring the Impact of Publications", in D. Tomov and L. Dimitrova (eds.), *Proceedings of the First National Conference with International Participation in Scientometrics and Linguistics of Scientific Text*, Varna.

Schubert, A., W. Glänzel and T. Braun (1987), "A New Methodology for Ranking Scientific Institutions", *Scientometrics*, **12**: 267-292.

Schubert, A., W. Glänzel, and T. Braun (1988), "Against Absolute Methods: Relative Scientometric Indicators and Relational Charts as Evaluation Tools", in A. F. J. van Raan (ed.), *Handbook of Quantitative Studies of Science and Technology*: 137-176.

Schubert, A., & Braun, C. (1986) Relative Indicators and Relational Charts for Comparative Assessment of Publication Output and Citation Impact. *Scientometrics*, 9, 281-291.

Schubert, A., & Braun, T. (1996) Cross-field Normalization of Scientometric Indicators. *Scientometrics*, 36, 311-324.

Schubert, A., Glänzel, W., Braun, T. (1983) Relative Citation Rate: A New Indicator for Measuring the Impact of Publications. in D. Tomov and L. Dimitrova (eds.), Proceedings of the First National Conference with International Participation in Scientometrics and Linguistics of Scientific Text, Varna.

Small, and Sweeney (1985), "Clustering the Science Citation Index Using Co-citations. I. A Comparison of Methods", *Scientometrics*, **7**: 393-404.

Vinkler, P. (1986), "Evaluation of Some Methods For the Relative Assessment of Scientific Publications", *Scientometrics*, **10**: 157-177.

Vinkler, P. (2003), "Relations of Relative Scientometric Indicators", *Scientometrics*, **58**: 687-694.

Van Eck, N.J., Waltman, L., Van Raan, A.F.J., Klautz, R.J.M., & Peul, W.C. (2012), "Citation analysis may severely underestimate the impact of clinical research as compared to basic research", Centre for Science and Technology Studies, Leiden University (arXiv:1210.0442).

Waltman, L, and N. J. van Eck (2012), "Universality of citation distributions revisited", *Journal of the American Society for Information Science and Technology*, **63**:72-77.

Waltman, L., & Van Eck, N. J.(2012), "Source normalized indicators of citation impact: An overview of different approaches and an empirical comparison", *Scientometrics*. arXiv:1208.6122

Waltman, L., and van Eck, N. J. (2013), "A systematic empirical comparison of different approaches for normalizing citation impact indicators", mimeo, Centre for Science and Technology Studies, Leiden University (arXiv:1301.4941).

Zitt M, and Bassecoulard, E. (1994), "Development of a Method for Detection and Trend Analysis of Research Fronts Built by Lexical or Cocitation Analysis", *Scientometrics*, **30**: 333-351.

Zitt M, and Small H. (2008), "Modifying the journal impact factor by fractional citation weighting: The audience factor", *Journal of the American Society for Information Science and Technology*, **59**: 1856-1860.

STATISTICAL APPENDIX

Table A1. Number of Articles and Mean Citation Rates by Field

	Number of Articles	%	Mean Citation	Standard Deviation
A. LIFE SCIENCES	**1,806,398**	**40.4**		
1. Biology & Biochemistry	275,568	6.2	12.6	20.1
2. Clinical Medicine	947,261	21.2	9.7	21.6
3. Immunology	60,875	1.4	16.0	23.0
4. Microbiology	73,039	1.6	11.4	13.9
5. Molecular Biology & Genetics	122,233	2.7	20.4	32.7
6. Neuroscience & Behav. Science	140,686	3.2	13.7	18.2
7. Pharmacology & Toxicology	76,728	1.7	8.0	11.0
8. Psychiatry & Psychology	110,008	2.5	7.0	11.3
B. PHYSICAL SCIENCES	**1,282,919**	**28.7**		
9. Chemistry	550,147	12.3	7.6	14.2
10. Computer Science	98,727	2.2	3.0	13.8
11. Mathematics	117,496	2.6	2.5	5.2
12. Physics	456,144	10.2	6.9	14.9
13. Space Science	60,405	1.4	11.0	20.5
C. OTHER NATURAL SCIENCES	**1,150,428**	**25.7**		
14. Agricultural Sciences	82,837	1.9	4.9	7.2
15. Engineering	356,269	8.0	3.2	5.8
16. Environment & Ecology	109,826	2.5	7.1	10.3
17. Geoscience	120,059	2.7	6.7	10.0
18. Materials Science	199,364	4.5	4.5	8.9
19. Multidisciplinary	20,672	0.5	3.2	7.0
20. Plant & Animal Science	261,401	5.8	5.1	8.0
D. SOCIAL SCIENCES	**232,587**	**5.2**		
21. Economics & Business	63,380	1.4	4.0	7.1
22. Social Sciences, General	169,207	3.8	3.3	5.7
ALL FIELDS	**4,472,332**	**100**	**7.9**	**16.4**

Table B1. Exchange Rates, Standard Deviations, and Coefficient of variation for the [356, 705] Interval

	Exchange Rates	Standard Deviation	Coefficient of Variation
	(1)	(2)	(3)
A. LIFE SCIENCES			
1. Biology & Biochemistry	18.1	1.0	0.053
2. Clinical Medicine	11.3	0.6	0.054
3. Immunology	23.8	1.9	0.078
4. Microbiology	18.1	1.4	0.079
5. Molecular Biology & Genetics	25.6	1.0	0.040
6. Neuroscience & Behav. Science	20.5	1.5	0.075
7. Pharmacology & Toxicology	11.6	0.9	0.078
8. Psychiatry & Psychology	8.8	0.8	0.091
B. PHYSICAL SCIENCES			
9. Chemistry	10.2	0.8	0.079
10. Computer Science	2.2	1.1	0.506
11. Mathematics	3.0	0.7	0.237
12. Physics	7.6	0.7	0.088
13. Space Science	13.7	1.0	0.072
C. OTHER NATURAL SCIENCES			
14. Agricultural Sciences	6.4	0.9	0.147
15. Engineering	3.7	0.6	0.167
16. Environment & Ecology	10.6	0.8	0.076
17. Geoscience	9.5	0.9	0.092
18. Materials Science	4.9	0.9	0.174
19. Multidisciplinary	2.4	1.1	0.472
20. Plant & Animal Science	7.0	0.6	0.092
D. SOCIAL SCIENCES			
21. Economics & Business	4.4	0.7	0.169
22. Social Sciences, General	3.9	0.6	0.165

Table A2. Number of Articles and Mean Citation Rates in the 219 Sub-fields and the 19 Fields in the Fractional Case

	Number of Articles	%	Mean Citation	Standard Deviation
	(1)	(2)	(3)	(4)
A. LIFE SCIENCES				
I. BIOSCIENCES	**342,480.5**	**7.67**	**15.8**	**20.1**
1. BIOLOGY	19,590.7	0.44	7.3	8.4
2. BIOLOGY, MISCELLANEOUS	277.1	0.01	3.3	0.9
3. EVOLUTIONARY BIOLOGY	5,953.0	0.13	12.6	11.5
4. BIOCHEMICAL RESEARCH METHODS	17,636.6	0.39	9.6	10.7
5. BIOCHEMISTRY & MOLECULAR BIOLOGY	161,192.8	3.61	17.4	19.7
6. BIOPHYSICS	28,162.4	0.63	10.9	8.3
7. CELL BIOLOGY	53,873.7	1.21	21.2	20.3
8. GENETICS & HEREDITY	43,311.1	0.97	15.8	20.3
9. DEVELOPMENTAL BIOLOGY	12,483.3	0.28	20.0	17.6
II. BIOMEDICAL RESEARCH	**247,383.6**	**5.54**	**9.0**	**9.9**
10. PATHOLOGY	22,487.5	0.50	9.9	11.7
11. ANATOMY & MORPHOLOGY	4,835.0	0.11	5.5	5.2
12. ENGINEERING, BIOMEDICAL	12,047.9	0.27	7.1	4.8
13. BIOTECHNOLOGY & APPLIED MICROBIOLOGY	37,682.5	0.84	9.2	11.4
14. MEDICAL LABORATORY TECHNOLOGY	8,619.5	0.19	6.6	8.9
15. MICROSCOPY	3,376.8	0.08	6.3	6.4
16. PHARMACOLOGY & PHARMACY	77,316.8	1.73	8.5	8.8
17. TOXICOLOGY	19,485.3	0.44	7.3	5.8
18. PHYSIOLOGY	29,551.8	0.66	10.9	7.9
19. MEDICINE, RESEARCH & EXPERIMENTAL	31,980.5	0.72	12.2	18.0
III. CLINICAL MEDICINE I (INTERNAL)	**440,082.7**	**9.86**	**12.6**	**22.8**
20. CARDIAC & CARDIOVASCULAR SYSTEMS	44591.9	1.00	10.2	12.3
21. RESPIRATORY SYSTEM	19873.3	0.45	10.1	8.9
22. ENDOCRINOLOGY & METABOLISM	47015.3	1.05	13.8	17.2
23. ANESTHESIOLOGY	16604.1	0.37	6.8	7.9
24. CRITICAL CARE MEDICINE	9488.3	0.21	11.5	11.4
25. EMERGENCY MEDICINE	5752.0	0.13	4.7	5.6

26. GASTROENTEROLOGY & HEPATOLOGY	35192.5	0.79	11.1	16.3
27. MEDICINE, GENERAL & INTERNAL	68428.2	1.53	13.6	51.5
28. TROPICAL MEDICINE	3793.3	0.08	5.4	3.4
29. HEMATOLOGY	33278.8	0.75	15.9	17.0
30. ONCOLOGY	74461.9	1.67	15.0	22.6
31. ALLERGY	5783.1	0.13	8.3	6.3
32. IMMUNOLOGY	53757.7	1.20	16.7	18.9
33. INFECTIOUS DISEASES	22062.3	0.49	11.3	9.2
IV. CLINICAL MEDICINE II (NON-INTERNAL)	**490,198.0**	**10.98**	**7.8**	**9.2**
34. GERIATRICS & GERONTOLOGY	6,566.1	0.15	7.9	6.2
35. OBSTETRICS & GYNECOLOGY	27,665.7	0.62	6.6	6.9
36. ANDROLOGY	1,663.5	0.04	5.7	6.8
37. REPRODUCTIVE BIOLOGY	10,972.9	0.25	10.2	7.6
38. GERONTOLOGY	4,473.6	0.10	6.8	5.1
39. DENTISTRY & ORAL SURGERY	22,405.0	0.50	5.3	6.1
40. DERMATOLOGY	21,692.7	0.49	6.2	8.1
41. UROLOGY & NEPHROLOGY	36,395.5	0.82	9.4	13.7
42. OTORHINOLARYNGOLOGY	16,012.2	0.36	4.0	3.7
43. OPHTHALMOLOGY	28,190.0	0.63	7.2	10.2
44. INTEGRATIVE & COMPLEMENTARY MEDICINE	1,708.3	0.04	4.2	4.0
45. CLINICAL NEUROLOGY	46,788.9	1.05	9.7	10.2
46. PSYCHIATRY	29,982.2	0.67	10.3	11.3
47. RADIOLOGY, NUCLEAR MED. & MED. IMAGING	45,722.9	1.02	8.0	9.5
48. ORTHOPEDICS	17,814.0	0.40	5.7	5.0
49. RHEUMATOLOGY	12,684.5	0.28	11.3	16.6
50. SPORT SCIENCES	15,515.9	0.35	5.8	5.4
51. SURGERY	74,364.1	1.67	6.4	6.5
52. TRANSPLANTATION	9,570.3	0.21	7.0	4.2
53. PERIPHERAL VASCULAR DISEASE	26,002.3	0.58	13.8	13.3
54. PEDIATRICS	34,007.5	0.76	6.1	7.7
V. CLINICAL MEDICINE III	**86,658.5**	**1.94**	**5.9**	**6.0**
55. HEALTH CARE SCIENCES & SERVICES	7,940.6	0.18	5.7	4.1
56. HEALTH POLICY & SERVICES	4,799.4	0.11	5.9	4.1
57. MEDICINE, LEGAL	3,991.6	0.09	4.4	5.1
58. NURSING	9,202.2	0.21	3.1	3.6
59. PUBLIC, ENV. & OCCUPATIONAL HEALTH	37,040.0	0.83	7.7	7.8

60. REHABILITATION	10,015.6	0.22	4.1	3.5
61. SUBSTANCE ABUSE	6,574.7	0.15	7.5	6.6
62. EDUCATION, SCIENTIFIC DISCIPLINES	4,667.8	0.10	2.9	2.3
63. MEDICAL INFORMATICS	2,426.8	0.05	4.1	2.1
VI. NEUROSCIENCES & BEHAVIORAL	**184,618.5**	**4.13**	**9.8**	**10.1**
64. NEUROIMAGING	2,603.3	0.06	10.8	5.6
65. NEUROSCIENCES	89,408.4	2.00	14.2	15.6
66. BEHAVIORAL SCIENCES	7,069.2	0.16	9.2	4.1
67. PSYCHOLOGY, BIOLOGICAL	1,760.5	0.04	7.5	3.4
68. PSYCHOLOGY	7,229.1	0.16	7.9	3.9
69. PSYCHOLOGY, APPLIED	6,307.8	0.14	5.0	5.0
70. PSYCHOLOGY, CLINICAL	14,166.8	0.32	7.1	6.9
71. PSYCHOLOGY, DEVELOPMENTAL	7,866.2	0.18	7.4	6.7
72. PSYCHOLOGY, EDUCATIONAL	4,820.3	0.11	4.8	5.3
73. PSYCHOLOGY, EXPERIMENTAL	11,416.3	0.26	7.0	6.2
74. PSYCHOLOGY, MATHEMATICAL	910.0	0.02	5.6	3.9
75. PSYCHOLOGY, MULTIDISCIPLINARY	16,339.0	0.37	4.3	7.7
76. PSYCHOLOGY, PSYCHOANALYSIS	2,109.6	0.05	2.2	2.9
77. PSYCHOLOGY, SOCIAL	9,586.7	0.21	6.6	8.4
78. SOCIAL SCIENCES, BIOMEDICAL	3,025.5	0.07	5.6	3.5
B. PHYSICAL SCIENCES				
VII. CHEMISTRY	**513,159.1**	**11.49**	**7.4**	**8.7**
79. CHEMISTRY, MULTIDISCIPLINARY	99,218.4	2.22	9.3	14.7
80. CHEMISTRY, INORGANIC & NUCLEAR	42,292.0	0.95	6.9	7.2
81. CHEMISTRY, ANALYTICAL	51,764.0	1.16	7.8	8.7
82. CHEMISTRY, APPLIED	17,483.2	0.39	4.8	2.8
83. ENGINEERING, CHEMICAL	44,458.1	1.00	4.1	4.2
84. CHEMISTRY, MEDICINAL	14,015.7	0.31	8.9	7.6
85. CHEMISTRY, ORGANIC	76,098.6	1.70	8.1	8.9
86. CHEMISTRY, PHYSICAL	95,580.2	2.14	8.0	7.9
87. ELECTROCHEMISTRY	15,409.6	0.35	7.1	6.2
88. POLYMER SCIENCE	56,839.4	1.27	6.5	8.8
VIII. PHYSICS	**522,921.8**	**11.71**	**6.4**	**11.2**
89. PHYSICS, MULTIDISCIPLINARY	92,884.0	2.08	8.5	20.2
90. SPECTROSCOPY	19,435.0	0.44	5.5	4.6

91. ACOUSTICS	10,604.0	0.24	4.1	3.8
92. OPTICS	45,132.7	1.01	5.4	6.9
93. PHYSICS, APPLIED	100,099.9	2.24	6.6	9.2
94. PHYSICS, ATOMIC, MOLECULAR & CHEMICAL	43,633.8	0.98	9.3	8.2
95. THERMODYNAMICS	7,968.4	0.18	3.4	1.8
96. PHYSICS, MATHEMATICAL	22,179.4	0.50	5.7	5.3
97. PHYSICS, NUCLEAR	18,519.7	0.41	5.7	7.4
98. PHYSICS, PARTICLES & SUB-FIELDS	28,648.3	0.64	10.1	20.6
99. PHYSICS, CONDENSED MATTER	86,321.6	1.93	6.3	8.6
100. PHYSICS OF SOLIDS, FLUIDS & PLASMAS	17,900.6	0.40	6.9	5.8
101. CRYSTALLOGRAPHY	29,594.6	0.66	4.0	28.9
IX. SPACE SCIENCES	**61,173.1**	**1.37**	**12.0**	**19.2**
102. ASTRONOMY & ASTROPHYSICS	61,173.1	1.37	12.0	19.2
X. MATHEMATICS	**139,956.3**	**3.13**	**2.8**	**9.4**
103. MATHEMATICS, APPLIED	41,617.9	0.93	2.7	3.2
104. STATISTICS & PROBABILITY	19,012.8	0.43	3.6	7.7
105. MATH., INTERDISCIPLINARY APPLICATIONS	8,159.0	0.18	4.1	2.6
106. SOCIAL SCIENCES, MATHEMATICAL METHODS	2,598.8	0.06	4.2	3.1
107. PURE MATHEMATICS	68,567.8	1.54	2.0	2.9
XI. COMPUTER SCIENCE	**113,370.0**	**2.54**	**3.4**	**5.8**
108. COMP. SCIENCE, ARTIFICIAL INTELLIGENCE	21,725.7	0.49	3.2	5.0
109. COMPUTER SCIENCE, CYBERNETICS	2,965.5	0.07	2.4	2.7
110. COMP. SCIENCE, HARDWARE & ARCHITECTURE	6,329.8	0.14	2.7	2.4
111. COMPUTER SCIENCE, INFORMATION SYSTEMS	12,870.5	0.29	3.1	3.6
112. COMP. SC., INTERDISCIPLINARY APPLICATIONS	13,659.9	0.31	4.2	5.3
113. COMP. SCIENCE, SOFTWARE ENGINEERING	12,780.8	0.29	2.7	3.3
114. COMPUTER SCIENCE, THEORY & METHODS	39,914.7	0.89	1.8	3.3
115. MATHEMATICAL & COMPUTATIONAL BIOLOGY	3,123.1	0.07	8.1	9.7
C. OTHER NATURAL SCIENCES				
XII. ENGINEERING	**288,058.5**	**6.45**	**3.3**	**3.4**
116. ENGINEERING, ELECTRICAL & ELECTRONIC	83,565.7	1.87	3.5	4.3
117. TELECOMMUNICATIONS	12,247.1	0.27	2.7	3.2
118. CONSTRUCTION & BUILDING TECHNOLOGY	4,639.8	0.10	2.5	1.7
119. ENGINEERING, CIVIL	12,516.2	0.28	2.2	1.8

120. ENGINEERING, ENVIRONMENTAL	9,672.1	0.22	7.1	5.0
121. ENGINEERING, MARINE	357.0	0.01	1.1	0.7
122. TRANSPORTATION SCIENCE & TECHNOLOGY	3,547.8	0.08	1.3	1.2
123. ENGINEERING, INDUSTRIAL	6,285.9	0.14	2.2	1.3
124. ENGINEERING, MANUFACTURING	6,932.4	0.16	2.4	1.5
125. ENGINEERING, MECHANICAL	26,333.2	0.59	2.6	2.4
126. MECHANICS	27,838.5	0.62	3.9	3.4
127. ROBOTICS	2,104.7	0.05	2.4	2.3
128. INSTRUMENTS & INSTRUMENTATION	17,583.1	0.39	3.5	2.2
129. IMAGING SCIENCE & PHOTOGR. TECHNOLOGY	2,679.8	0.06	4.3	3.1
130. ENERGY & FUELS	12,929.4	0.29	3.7	3.0
131. NUCLEAR SCIENCE & TECHNOLOGY	21,161.0	0.47	2.8	2.6
132. ENGINEERING, PETROLEUM	3,566.8	0.08	1.0	1.1
133. AUTOMATION & CONTROL SYSTEMS	9,343.5	0.21	2.8	2.7
134. ENGINEERING, MULTIDISCIPLINARY	11,279.3	0.25	2.6	2.2
135. ERGONOMICS	1,382.3	0.03	3.2	1.5
136. OPERATIONS RES. & MANAGEMENT SCIENCE	12,092.9	0.27	2.9	2.6
XIII. MATERIALS SCIENCE	**185,225.7**	**4.15**	**4.4**	**5.1**
137. MATERIALS SCIENCE, MULTIDISCIPLINARY	90,734.1	2.03	4.5	4.7
138. MATERIALS SCIENCE, BIOMATERIALS	3,953.5	0.09	10.2	5.8
139. MATERIALS SCIENCE, CERAMICS	18,866.3	0.42	3.5	4.8
140. MAT. SC., CHARACTERIZATION & TESTING	5,159.8	0.12	1.4	2.4
141. MATERIALS SCIENCE, COATINGS & FILMS	10,519.9	0.24	5.6	3.3
142. MATERIALS SCIENCE, COMPOSITES	7,957.8	0.18	2.9	3.9
143. MATERIALS SCIENCE, PAPER & WOOD	6,000.6	0.13	1.8	2.4
144. MATERIALS SCIENCE, TEXTILES	3,656.8	0.08	1.8	2.0
145. METALL. & METALLURGICAL ENGINEERING	29,468.1	0.66	2.8	3.3
146. NANOSCIENCE & NANOTECHNOLOGY	8,908.6	0.20	6.1	4.1
XIV. GEOSCIENCES	**144,907.0**	**3.25**	**6.0**	**7.0**
147. GEOCHEMISTRY & GEOPHYSICS	27,878.1	0.62	7.4	10.4
148. GEOGRAPHY, PHYSICAL	4,368.3	0.10	7.0	3.8
149. GEOLOGY	7,291.2	0.16	6.5	7.3
150. ENGINEERING, GEOLOGICAL	2,717.6	0.06	2.8	1.8
151. PALEONTOLOGY	5,862.2	0.13	3.9	3.5
152. REMOTE SENSING	2,389.6	0.05	5.6	3.4
153. OCEANOGRAPHY	13,918.8	0.31	7.6	6.6

154. ENGINEERING, OCEAN	1,928.3	0.04	2.6	2.6
155. METEOROLOGY & ATMOSPHERIC SCIENCES	23,267.3	0.52	9.2	11.0
156. ENGINEERING, AEROSPACE	10,028.8	0.22	1.8	2.4
157. MINERALOGY	5,410.5	0.12	5.3	4.8
158. MINING & MINERAL PROCESSING	3,672.2	0.08	2.4	1.9
159. GEOSCIENCES, MULTIDISCIPLINARY	36,174.3	0.81	5.5	5.9
XV. AGRICULTURAL & ENVIRONMENT	**180,472.2**	**4.04**	**5.6**	**6.1**
160. AGRICULTURAL ENGINEERING	3,675.5	0.08	3.2	2.9
161. AGRICULTURE, MULTIDISCIPLINARY	11,518.7	0.26	3.5	3.3
162. AGRONOMY	16,837.2	0.38	3.8	3.5
163. LIMNOLOGY	2,742.4	0.06	7.3	3.8
164. SOIL SCIENCE	11,948.1	0.27	5.4	5.7
165. BIODIVERSITY CONSERVATION	3,507.3	0.08	5.6	3.3
166. ENVIRONMENTAL SCIENCES	44,640.7	1.00	6.6	5.4
167. ENVIRONMENTAL STUDIES	5,592.3	0.13	3.5	2.3
168. FOOD SCIENCE & TECHNOLOGY	31,783.8	0.71	4.7	3.9
169. NUTRITION & DIETETICS	19,574.3	0.44	9.2	10.8
170. AGRICULTURE, DAIRY & ANIMAL SCIENCE	20,968.0	0.47	3.6	4.4
171. HORTICULTURE	7,683.9	0.17	3.3	2.6
XVI. BIOLOGY (ORGANISMIC AND SUPRAORGONISMIC LEVEL)	**323,550.6**	**7.25**	**7.0**	**8.0**
172. ORNITHOLOGY	5,141.0	0.12	4.2	7.7
173. ZOOLOGY	28,223.6	0.63	4.9	4.5
174. ENTOMOLOGY	20,111.8	0.45	3.6	4.0
175. WATER RESOURCES	13,317.7	0.30	4.4	2.8
176. FISHERIES	12,410.6	0.28	4.7	3.5
177. MARINE & FRESHWATER BIOLOGY	23,026.3	0.52	5.7	3.9
178. MICROBIOLOGY	44,835.5	1.00	11.0	9.8
179. PARASITOLOGY	9,784.2	0.22	6.1	6.3
180. VIROLOGY	19,375.5	0.43	15.1	14.8
181. FORESTRY	10,665.6	0.24	5.2	5.5
182. MYCOLOGY	5,700.2	0.13	4.3	5.4
183. PLANT SCIENCES	53,680.8	1.20	7.4	9.0
184. ECOLOGY	28,265.6	0.63	8.6	7.3
185. VETERINARY SCIENCES	49,012.4	1.10	3.2	4.0

XVII. MULTIDISCIPLINARY	**27,218.9**	**0.61**	**3.2**	**6.5**
186. MULTIDISCIPLINARY SCIENCES	27,218.9	0.61	3.2	6.5
D. SOCIAL SCIENCES				
XVIII. SOCIAL SCIENCES, GENERAL	**118,297.3**	**2.65**	**3.0**	**3.6**
187. CRIMINOLOGY & PENOLOGY	2,777.0	0.06	3.5	4.2
188. LAW	8,529.8	0.19	3.5	4.7
189. POLITICAL SCIENCE	10,838.3	0.24	2.5	4.1
190. PUBLIC ADMINISTRATION	3,036.5	0.07	2.6	3.1
191. ETHNIC STUDIES	701.3	0.02	1.7	1.1
192. FAMILY STUDIES	3,166.8	0.07	4.0	3.0
193. SOCIAL ISSUES	2,771.7	0.06	2.6	3.2
194. SOCIAL WORK	3,880.8	0.09	2.4	2.2
195. SOCIOLOGY	10,554.0	0.24	3.0	4.7
196. WOMEN'S STUDIES	2,656.7	0.06	2.4	2.3
197. EDUCATION & EDUCATIONAL RESEARCH	14,580.3	0.33	2.2	3.0
198. EDUCATION, SPECIAL	2,076.2	0.05	3.4	2.7
199. AREA STUDIES	3,197.6	0.07	1.3	1.8
200. GEOGRAPHY	4,487.6	0.10	4.3	4.9
201. PLANNING & DEVELOPMENT	4,041.8	0.09	3.2	2.9
202. TRANSPORTATION	1,050.8	0.02	3.0	1.7
203. URBAN STUDIES	2,802.9	0.06	3.1	2.4
204. ETHICS	2,208.6	0.05	2.1	1.6
205. MEDICAL ETHICS	305.3	0.01	3.8	1.2
206. ANTHROPOLOGY	5,620.2	0.13	2.7	3.2
207. COMMUNICATION	4,085.0	0.09	3.1	3.2
208. DEMOGRAPHY	1,749.8	0.04	4.2	4.9
209. HISTORY OF SOCIAL SCIENCES	867.0	0.02	1.3	1.0
210. INFORMATION SCIENCE & LIBRARY SCIENCE	7,034.7	0.16	2.4	2.9
211. INTERNATIONAL RELATIONS	4,820.8	0.11	2.3	3.6
212. LINGUISTICS	3,921.7	0.09	3.8	3.0
213. SOCIAL SCIENCES, INTERDISCIPLINARY	6,534.3	0.15	2.3	2.5
XIX. ECONOMICS & BUSINESS	**55,615.8**	**1.25**	**4.1**	**5.1**
214. AGRICULTURAL ECONOMICS & POLICY	1,005.5	0.02	2.8	1.8
215. ECONOMICS	30,439.6	0.68	3.5	5.2
216. INDUSTRIAL RELATIONS & LABOR	1,917.7	0.04	3.0	3.5
217. BUSINESS	7,255.2	0.16	5.0	5.1

218. BUSINESS, FINANCE	5,351.8	0.12		4.9	6.7
219. MANAGEMENT	9,646.2	0.22		4.5	4.3
Total	**4,465,348**	**100.00**	**Mean**	**5.9**	**3.6**
			Std	**6.4**	**5.6**

Table 1A. Citation Inequality Decomposition for Different Inequality Indices and Different Quantile Choices

Inequality Indices	Quantile Choice, Π	Within-group Term, W	Skew. of Sc. Term, S	*ICP* Term	Total Citation Ineq., $I_a(C)$	Percentages In %: (1)/(4)	(2)/(4)	(3)/(4)
		(1)	(2)	(3)	(4)	(5)	(6)	(7)
$a = 2$	10	1.0407	0.8547	0.2866	2.1820	47.7	39.2	13.13
	50	0.6817	1.1437	0.3570	2.1820	31.2	52.4	16.36
	100	0.5771	1.2275	0.3778	2.1820	26.4	56.2	17.31
	1,000	0.3072	1.4415	0.4334	2.1820	14.1	66.1	19.86
$a = 0$, $\varepsilon_1 = 0.1$	10	0.0702	0.8905	0.2093	1.1700	6.0	76.1	17.89
	50	0.0134	0.9237	0.2329	1.1700	1.2	79.0	19.90
	100	0.0063	0.9306	0.2331	1.1700	0.5	79.5	19.92
	1,000	0.0007	0.9273	0.2419	1.1700	0.1	79.3	20.68
$a = 0$, $\varepsilon_2 = 0.01$	10	0.1644	1.0168	0.4445	1.6258	10.1	62.5	27.34
	50	0.0316	1.1049	0.4893	1.6258	1.9	68.0	30.10
	100	0.0170	1.1093	0.4995	1.6258	1.0	68.2	30.72
	1,000	0.0017	1.1111	0.5129	1.6258	0.1	68.3	31.55
$a = 1$, $\varepsilon_1 = 0.1$	10	0.0914	0.6439	0.1120	0.8473	10.8	76.0	13.22
	50	0.0293	0.7024	0.1150	0.8473	3.5	83.0	13.58
	100	0.0188	0.7124	0.1154	0.8473	2.2	84.1	13.62
	1,000	0.0045	0.7265	0.1161	0.8473	0.5	85.8	13.70
$a = 1$, $\varepsilon_2 = 0.01$	10	0.0937	0.6613	0.1168	0.8721	10.7	75.9	13.40
	50	0.0296	0.7226	0.1202	0.8721	3.4	82.8	13.79
	100	0.0191	0.7328	0.1206	0.8721	2.2	84.0	13.83
	1,000	0.0046	0.7462	0.1211	0.8721	0.5	85.6	13.89
$a = 1$, ln(0)=0	10	0.0940	0.6636	0.1179	0.8755	10.7	75.8	13.46
	50	0.0300	0.7244	0.1211	0.8755	3.4	87.2	13.83
	100	0.0192	0.7348	0.1215	0.8755	2.2	83.9	13.88
	1,000	0.0046	0.7488	0.1221	0.8755	0.52	85.53	13.95

Table 1B. Citation Inequality Decomposition at the Sub-field Level

A. FRACTIONAL CASE	Within-group Term, ***W*** (1)	Skewness of Science Term, ***S*** (2)	***IDCP*** Term (3)	Overall Inequality (4)	Percentages In %: (1)/(4)	(2)/(4)	(3)/(4)
	0.0030	0.7062	0.1552	0.8644	0.35	81.70	17.95

B. MULTIPLICATIVE CASE	***W'*** (1)	***S'*** (2)	***IDCP'*** (3)	Overall Inequality (4)	Percentages In %: (1)/(4)	(2)/(4)	(3)/(4)
	0.0030	0.6950	0.1544	0.8524	0.35	81.54	18.11

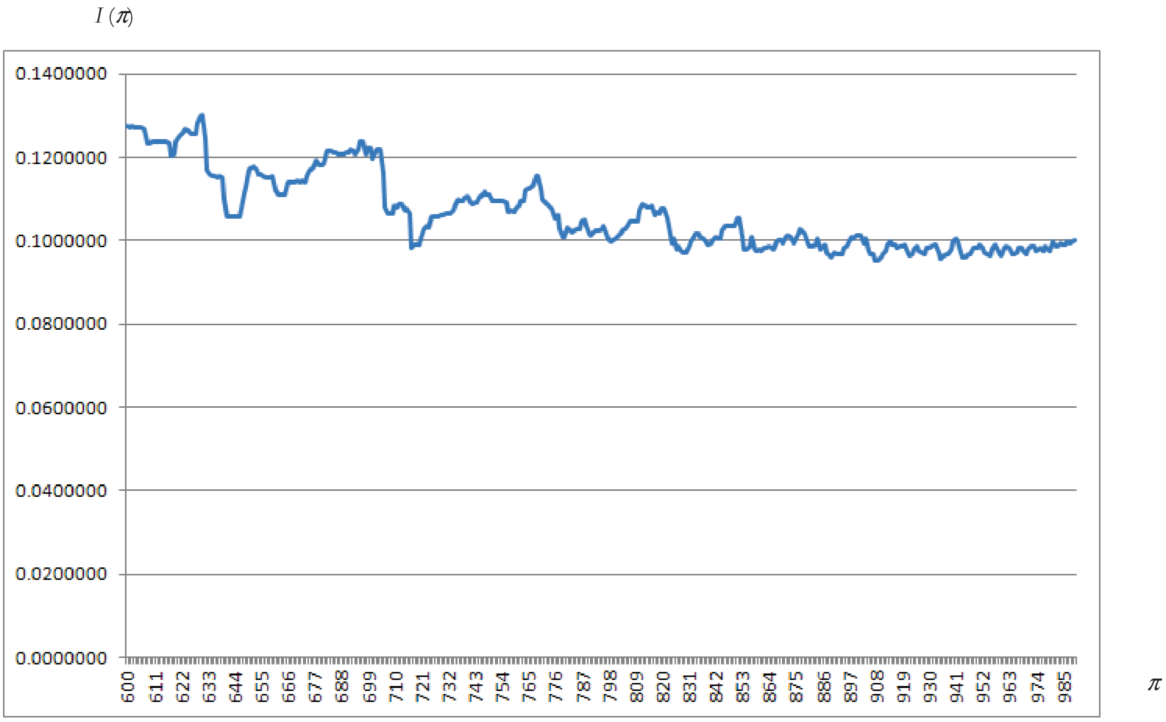


Figure 1A. Citation Inequality Due to Differences in Citation Practices, $I(\pi)$ versus π. Raw Data.

Table 2A. Exchange Rates, Standard Deviations, and Coefficient of variation for the [706, 998] Interval, and Exchange Rates Based on Mean Citations

	Exchange Rates	Standard Deviation	Coefficient of Variation	% of Citations	ERs Based on Mean Citations	ERs Based on Mean Cits. In the [706, 998] Interval
	(1)	(2)	(3)	(4)	(5)	(6)
1. Biology & Biochemistry	15.8	0.9	0.054	68.0	16.0	15.3
2. Clinical Medicine	12.1	0.6	0.049	71.8	12.4	12.5
3. Immunology	19.5	0.9	0.048	66.3	20.4	19.0
4. Microbiology	14.4	1.3	0.092	65.8	14.6	13.5
5. Molecular Biology & Genetics	25.7	0.6	0.022	71.1	25.9	25.9
6. Neuroscience & Behav. Science	17.1	0.8	0.050	67.2	17.5	16.5
7. Pharmacology & Toxicology	10.1	0.6	0.056	68.4	10.2	9.8
8. Psychiatry & Psychology	9.1	0.2	0.025	72.4	9.0	9.1
9. Chemistry	9.9	0.4	0.037	70.9	9.7	9.7
10. Computer Science	3.7	0.5	0.124	76.3	3.8	4.0
11. Mathematics	3.3	0.2	0.059	75.4	3.1	3.3
12. Physics	8.8	0.5	0.061	74.2	8.7	9.1
13. Space Science	14.2	0.3	0.019	71.9	14.0	14.2
14. Agricultural Sciences	6.5	0.4	0.056	72.5	6.2	6.3
15. Engineering	4.4	0.2	0.054	75.9	4.1	4.4
16. Environment & Ecology	9.1	0.7	0.073	68.3	9.1	8.7
17. Geoscience	8.9	0.6	0.069	70.1	8.6	8.5
18. Materials Science	5.9	0.3	0.048	75.0	5.8	6.1
19. Multidisciplinary	4.3	0.7	0.158	81.6	4.1	4.7
20. Plant & Animal Science	6.7	0.3	0.045	71.3	6.5	6.5
21. Economics & Business	5.2	0.4	0.068	75.6	5.0	5.3
22. Social Sciences, General	4.5	0.2	0.045	75.1	4.2	4.5
Mean				**72.1**		

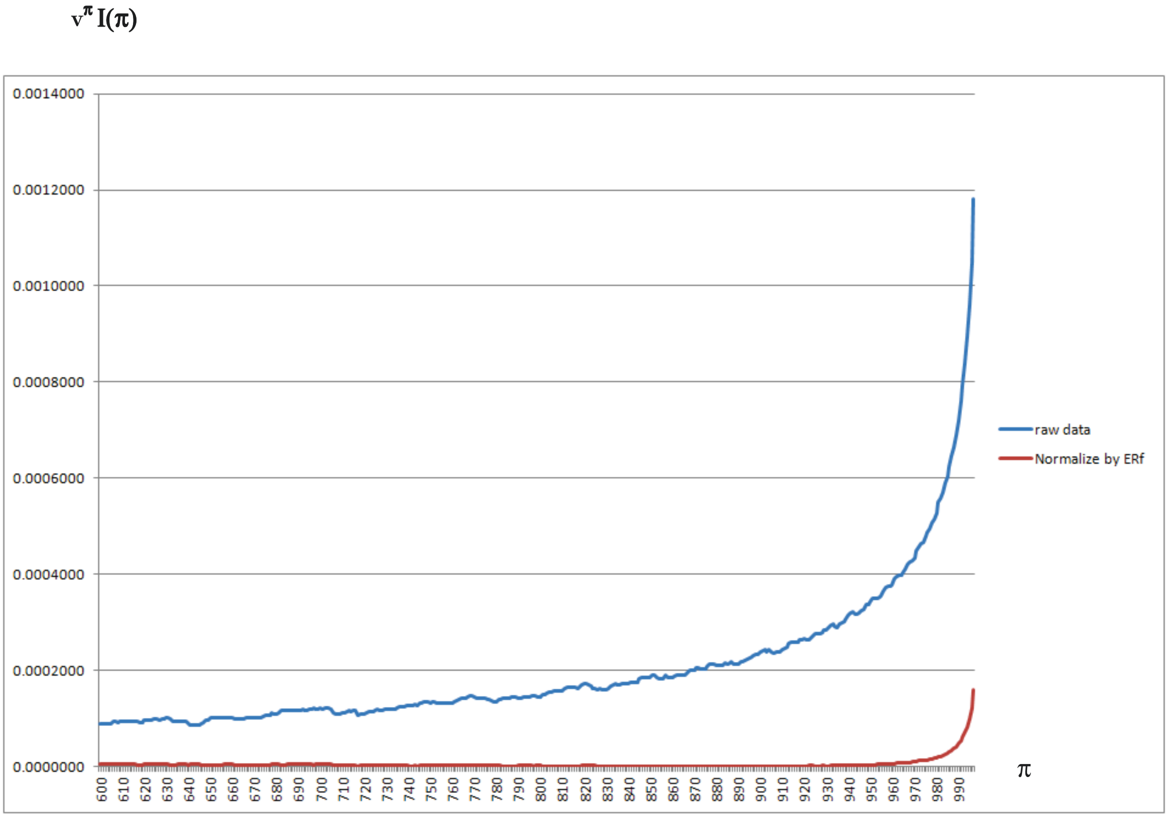


Figure 2A. Weighted Citation Inequality Due to Differences in Citation Practices, v^{π} I(π) vs. π. Raw *vs.* Normalized Data

Table 3A. Total Citation Inequality Decomposition Before and After Normalization: *IDCP* Interval Detail

	Within-group Term, W	Skew. of Sc. Term, S	*IDCP* Term	Total Citation Ineq., $I_1(C)$	Percentages In %:		
					(1)/(4)	(2)/(4)	(3)/(4)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
A. RAW DATA							
All Quantiles	0.0046	0.7488	0.1221	0.8755	0.53	85.52	13.95
[1, 705]			0.0449				5.13
[706, 998]			0.0717				8.18
[999, 1000]			0.0056				0.64
B. EXCHANGE RATE NORMALIZATION							
All Quantiles	0.0051	0.7788	0.0167	0.8006	0.63	97.28	2.09
[1, 705]			0.0127				1.59
[706, 998]			0.0018				0.23
[999, 1000]			0.0022				0.27
C. NORMALIZATION WITH TWO EXCHANGE RATES							
All Quantiles	0.0050	0.7715	0.0147	0.7913	0.64	97.50	1.86
[1, 705]			0.0108				1.36
[706, 998]			0.0018				0.23
[999, 1000]			0.0021				0.27
D. MEAN NORMALIZATION							
All Quantiles	0.0050	0.7794	0.0164	0.8008	0.63	97.32	2.05
[1, 705]			0.0124				1.55
[706, 998]			0.0020				0.25
[999, 1000]			0.0020				0.25
E. GLÄNZEL NORMALIZATION							
All Quantiles	0.0048	0.7638	0.0241	0.7928	0.61	96.35	3.05
[1, 705]			0.0184				2.32
[706, 998]			0.0047				0.60
[999, 1000]			0.0010				0.13

Table 2B. Exchange Rates, Standard Deviations, and Coefficients of Variation for the [661, 978] Interval In the Fractional Approach

		Exchange Rates	Standard Deviation	Coefficient of Variation	% of Citations	Exch. Rates Based on Mean Citations
		(1)	(2)	(3)	(4)	(5)
	A. LIFE SCIENCES					
	I. BIOSCIENCES					
1	BIOLOGY	10.3	0.3	0.032	64.1	9.8
2	BIOLOGY, MISCELLANEOUS	5.0	0.3	0.063	65.4	4.6
3	EVOLUTIONARY BIOLOGY	16.1	1.8	0.109	56.3	16.4
4	BIOCHEMICAL RESEARCH METHODS	11.5	0.7	0.060	52.9	12.8
5	BIOCHEMISTRY & MOLECULAR BIOLOGY	20.6	0.5	0.023	58.2	21.2
6	BIOPHYSICS	14.0	0.7	0.053	58.7	14.1
7	CELL BIOLOGY	26.9	0.9	0.032	60.3	27.3
8	GENETICS & HEREDITY	19.4	0.4	0.022	57.7	20.5
9	DEVELOPMENTAL BIOLOGY	23.4	0.4	0.016	59.0	24.0
	II. BIOMEDICAL RESEARCH					
10	PATHOLOGY	11.8	0.3	0.023	62.3	11.5
11	ANATOMY & MORPHOLOGY	7.7	0.5	0.066	60.9	7.4
12	ENGINEERING, BIOMEDICAL	9.5	0.5	0.053	61.3	9.1
13	BIOTECHNOLOGY & APPLIED MICROBIOLOGY	11.5	0.3	0.024	58.0	11.9
14	MEDICAL LABORATORY TECHNOLOGY	8.1	0.3	0.031	62.0	7.9
15	MICROSCOPY	8.6	0.7	0.077	60.8	8.3
16	PHARMACOLOGY & PHARMACY	10.6	0.5	0.046	60.0	10.5
17	TOXICOLOGY	9.7	0.7	0.071	58.9	9.6
18	PHYSIOLOGY	14.0	1.4	0.102	59.4	13.5
19	MEDICINE, RESEARCH & EXPERIMENTAL	15.4	2.6	0.171	61.2	16.5
	III. CLINICAL MEDICINE I (INTERNAL)					
20	CARDIAC & CARDIOVASCULAR SYSTEMS	14.9	1.0	0.070	61.6	15.1
21	RESPIRATORY SYSTEM	13.7	0.7	0.051	60.6	13.4
22	ENDOCRINOLOGY & METABOLISM	16.9	1.1	0.066	58.3	16.9
23	ANESTHESIOLOGY	9.2	0.3	0.037	62.8	8.8
24	CRITICAL CARE MEDICINE	14.8	0.5	0.036	61.9	14.2
25	EMERGENCY MEDICINE	5.8	0.3	0.050	62.8	5.5
26	GASTROENTEROLOGY & HEPATOLOGY	13.5	0.3	0.022	60.1	13.6
27	MEDICINE, GENERAL & INTERNAL	12.0	4.9	0.405	52.1	16.7
28	TROPICAL MEDICINE	7.2	0.5	0.074	62.1	6.8
29	HEMATOLOGY	22.2	0.3	0.014	60.2	22.3
30	ONCOLOGY	18.0	0.6	0.031	58.6	18.3
31	ALLERGY	12.2	0.5	0.038	63.1	11.5
32	IMMUNOLOGY	17.8	0.3	0.017	59.0	18.3
33	INFECTIOUS DISEASES	15.4	1.0	0.068	59.6	15.1

	IV. CLINICAL MEDICINE II (NON-INTERNAL)					
34	GERIATRICS & GERONTOLOGY	11.2	0.6	0.051	60.9	10.9
35	OBSTETRICS & GYNECOLOGY	9.2	0.4	0.044	62.3	8.8
36	ANDROLOGY	7.3	0.5	0.068	60.3	7.1
37	REPRODUCTIVE BIOLOGY	12.5	1.1	0.089	59.0	12.3
38	GERONTOLOGY	10.2	0.5	0.049	62.7	9.6
39	DENTISTRY & ORAL SURGERY	7.2	0.6	0.077	60.6	6.9
40	DERMATOLOGY	8.2	0.3	0.038	62.1	7.9
41	UROLOGY & NEPHROLOGY	12.3	0.3	0.025	61.6	12.0
42	OTORHINOLARYNGOLOGY	6.0	0.4	0.069	62.5	5.6
43	OPHTHALMOLOGY	9.5	0.3	0.034	61.7	9.2
44	INTEGRATIVE & COMPLEMENTARY MEDICINE	6.3	0.6	0.097	61.4	5.9
45	CLINICAL NEUROLOGY	12.4	0.3	0.023	61.3	12.1
46	PSYCHIATRY	13.1	0.3	0.019	62.0	12.7
47	RADIOLOGY, NUCLEAR MED. & MED. IMAGING	10.1	0.3	0.026	61.5	9.9
48	ORTHOPEDICS	7.9	0.3	0.043	61.6	7.6
49	RHEUMATOLOGY	14.6	0.6	0.041	59.7	14.5
50	SPORT SCIENCES	8.1	0.5	0.064	62.2	7.7
51	SURGERY	8.5	0.2	0.028	61.9	8.3
52	TRANSPLANTATION	9.5	0.2	0.026	61.9	9.2
53	PERIPHERAL VASCULAR DISEASE	20.2	0.3	0.013	59.8	20.4
54	PEDIATRICS	7.7	0.3	0.035	62.1	7.5
	V. CLINICAL MEDICINE III					
55	HEALTH CARE SCIENCES & SERVICES	7.9	0.5	0.061	60.3	7.7
56	HEALTH POLICY & SERVICES	8.4	0.4	0.042	59.3	8.5
57	MEDICINE, LEGAL	5.8	0.4	0.072	60.5	5.6
58	NURSING	4.3	0.4	0.090	61.9	4.1
59	PUBLIC, ENV. & OCCUPATIONAL HEALTH	9.7	0.3	0.034	60.8	9.5
60	REHABILITATION	5.9	0.4	0.065	62.2	5.6
61	SUBSTANCE ABUSE	9.8	0.9	0.096	59.2	9.6
62	EDUCATION, SCIENTIFIC DISCIPLINES	4.0	0.3	0.068	64.9	3.7
63	MEDICAL INFORMATICS	5.7	0.3	0.045	62.9	5.5
	VI. NEUROSCIENCES & BEHAVIORAL					
64	NEUROIMAGING	14.6	0.4	0.025	63.1	14.0
65	NEUROSCIENCES	16.9	0.5	0.031	59.6	16.9
66	BEHAVIORAL SCIENCES	11.5	1.4	0.119	56.0	11.7
67	PSYCHOLOGY, BIOLOGICAL	9.9	0.9	0.086	56.9	10.1
68	PSYCHOLOGY	10.3	0.7	0.068	60.6	9.9
69	PSYCHOLOGY, APPLIED	6.4	0.4	0.070	62.4	6.0
70	PSYCHOLOGY, CLINICAL	9.9	0.4	0.042	60.6	9.7
71	PSYCHOLOGY, DEVELOPMENTAL	10.6	0.5	0.051	60.8	10.2
72	PSYCHOLOGY, EDUCATIONAL	6.8	0.3	0.040	64.2	6.5
73	PSYCHOLOGY, EXPERIMENTAL	10.2	0.5	0.046	61.2	9.9
74	PSYCHOLOGY, MATHEMATICAL	6.9	0.3	0.038	61.3	6.8
75	PSYCHOLOGY, MULTIDISCIPLINARY	6.2	0.5	0.087	63.3	6.2
76	PSYCHOLOGY, PSYCHOANALYSIS	3.7	0.4	0.106	67.8	3.4

77	PSYCHOLOGY, SOCIAL	8.3	0.3	0.032	61.5	8.2
78	SOCIAL SCIENCES, BIOMEDICAL	7.2	0.3	0.047	61.2	7.0
	B. PHYSICAL SCIENCES					
	VII. CHEMISTRY					
79	CHEMISTRY, MULTIDISCIPLINARY	11.9	1.2	0.103	65.4	11.5
80	CHEMISTRY, INORGANIC & NUCLEAR	9.2	0.7	0.074	61.4	8.8
81	CHEMISTRY, ANALYTICAL	9.9	0.4	0.044	60.5	9.7
82	CHEMISTRY, APPLIED	7.6	0.5	0.070	62.3	7.2
83	ENGINEERING, CHEMICAL	6.0	0.3	0.044	63.7	5.7
84	CHEMISTRY, MEDICINAL	9.8	0.8	0.083	59.4	9.6
85	CHEMISTRY, ORGANIC	10.7	1.0	0.096	59.3	10.4
86	CHEMISTRY, PHYSICAL	10.5	0.5	0.047	60.5	10.3
87	ELECTROCHEMISTRY	10.2	0.8	0.076	60.4	9.9
88	POLYMER SCIENCE	8.2	0.3	0.031	61.4	8.1
	VIII. PHYSICS					
89	PHYSICS, MULTIDISCIPLINARY	10.0	1.7	0.169	61.8	10.5
90	SPECTROSCOPY	7.6	0.4	0.050	62.1	7.3
91	ACOUSTICS	5.5	0.3	0.055	63.3	5.2
92	OPTICS	7.3	0.3	0.036	62.7	7.0
93	PHYSICS, APPLIED	7.5	0.4	0.048	60.7	7.6
94	PHYSICS, ATOMIC, MOLECULAR & CHEMICAL	11.0	0.8	0.074	59.8	10.7
95	THERMODYNAMICS	4.8	0.4	0.080	61.6	4.6
96	PHYSICS, MATHEMATICAL	7.3	0.3	0.035	61.7	7.2
97	PHYSICS, NUCLEAR	6.2	0.4	0.065	62.0	6.2
98	PHYSICS, PARTICLES & SUB-FIELDS	10.8	1.1	0.102	59.8	11.4
99	PHYSICS, CONDENSED MATTER	7.4	0.3	0.045	61.4	7.4
100	PHYSICS OF SOLIDS, FLUIDS & PLASMAS	9.3	0.6	0.063	59.8	9.1
101	CRYSTALLOGRAPHY	5.1	0.3	0.053	58.8	5.2
	IX. SPACE SCIENCES					
102	ASTRONOMY & ASTROPHYSICS	14.8	0.3	0.018	60.6	14.8
	X. MATHEMATICS					
103	MATHEMATICS, APPLIED	3.9	0.2	0.062	65.7	3.6
104	STATISTICS & PROBABILITY	5.2	0.5	0.098	52.5	6.2
105	MATH., INTERDISCIPLINARY APPLICATIONS	5.6	0.3	0.045	60.8	5.6
106	SOCIAL SCIENCES, MATHEMATICAL METHODS	5.5	0.3	0.045	61.4	5.5
107	PURE MATHEMATICS	2.8	0.2	0.087	66.4	2.6
	XI. COMPUTER SCIENCE					
108	COMP. SCIENCE, ARTIFICIAL INTELLIGENCE	5.4	0.6	0.118	63.3	5.4
109	COMPUTER SCIENCE, CYBERNETICS	3.6	0.4	0.108	66.7	3.4
110	COMP. SCIENCE, HARDWARE & ARCHITECTURE	4.0	0.5	0.124	61.4	4.1
111	COMPUTER SCIENCE, INFORMATION SYSTEMS	4.4	0.6	0.143	62.4	4.5

112	COMP. SC., INTERDISCIPLINARY APPLICATIONS	5.5	0.6	0.102	58.1	6.0
113	COMP. SCIENCE, SOFTWARE ENGINEERING	3.6	0.4	0.107	65.5	3.4
114	COMPUTER SCIENCE, THEORY & METHODS	3.1	0.4	0.115	65.5	3.0
115	MATHEMATICAL & COMPUTATIONAL BIOLOGY	9.8	0.4	0.044	52.9	11.4
	C. OTHER NATURAL SCIENCES					
	XII. ENGINEERING					
116	ENGINEERING, ELECTRICAL & ELECTRONIC	4.7	0.4	0.077	63.1	4.6
117	TELECOMMUNICATIONS	3.8	0.5	0.144	62.2	3.9
118	CONSTRUCTION & BUILDING TECHNOLOGY	3.5	0.3	0.090	65.4	3.1
119	ENGINEERING, CIVIL	3.4	0.3	0.086	67.0	3.1
120	ENGINEERING, ENVIRONMENTAL	9.1	0.3	0.035	62.4	8.7
121	ENGINEERING, MARINE	1.6	0.3	0.212	71.5	1.4
122	TRANSPORTATION SCIENCE & TECHNOLOGY	2.1	0.5	0.227	69.9	2.0
123	ENGINEERING, INDUSTRIAL	3.3	0.3	0.091	66.6	2.9
124	ENGINEERING, MANUFACTURING	3.6	0.3	0.089	64.8	3.2
125	ENGINEERING, MECHANICAL	3.9	0.2	0.060	63.7	3.7
126	MECHANICS	5.2	0.3	0.050	63.8	4.9
127	ROBOTICS	3.8	0.2	0.065	65.0	3.6
128	INSTRUMENTS & INSTRUMENTATION	5.1	0.3	0.051	65.0	4.7
129	IMAGING SCIENCE & PHOTOGR. TECHNOLOGY	7.4	0.4	0.061	64.6	7.0
130	ENERGY & FUELS	5.0	0.3	0.064	64.9	4.7
131	NUCLEAR SCIENCE & TECHNOLOGY	4.4	0.3	0.061	64.0	4.1
132	ENGINEERING, PETROLEUM	1.7	0.4	0.255	73.5	1.5
133	AUTOMATION & CONTROL SYSTEMS	4.1	0.2	0.059	63.8	3.9
134	ENGINEERING, MULTIDISCIPLINARY	3.9	0.4	0.089	66.0	3.7
135	ERGONOMICS	4.8	0.4	0.088	63.0	4.4
136	OPERATIONS RES. & MANAGEMENT SCIENCE	4.1	0.2	0.060	63.6	3.8
	XIII. MATERIALS SCIENCE					
137	MATERIALS SCIENCE, MULTIDISCIPLINARY	6.4	0.4	0.056	60.7	6.4
138	MATERIALS SCIENCE, BIOMATERIALS	13.0	1.1	0.085	59.3	12.7
139	MATERIALS SCIENCE, CERAMICS	4.7	0.3	0.074	68.3	4.2
140	MAT. SC., CHARACTERIZATION & TESTING	2.2	0.4	0.167	70.6	2.0
141	MATERIALS SCIENCE, COATINGS & FILMS	7.5	0.4	0.057	61.0	7.3
142	MATERIALS SCIENCE, COMPOSITES	3.4	0.3	0.087	65.9	3.1
143	MATERIALS SCIENCE, PAPER & WOOD	2.9	0.3	0.092	68.1	2.6
144	MATERIALS SCIENCE, TEXTILES	2.9	0.3	0.095	65.5	2.7
145	METALL. & METALLURGICAL ENGINEERING	4.7	0.4	0.089	63.5	4.7
146	NANOSCIENCE & NANOTECHNOLOGY	8.0	0.3	0.036	60.0	8.1

	XIV. GEOSCIENCES					
147	GEOCHEMISTRY & GEOPHYSICS	9.7	0.6	0.066	61.5	9.3
148	GEOGRAPHY, PHYSICAL	9.1	0.9	0.097	59.8	8.8
149	GEOLOGY	8.0	0.5	0.061	62.4	7.5
150	ENGINEERING, GEOLOGICAL	3.8	0.3	0.093	62.1	3.6
151	PALEONTOLOGY	6.5	0.4	0.057	63.7	6.1
152	REMOTE SENSING	7.8	0.3	0.037	60.8	7.8
153	OCEANOGRAPHY	10.1	1.0	0.101	61.6	9.5
154	ENGINEERING, OCEAN	3.6	0.4	0.106	66.7	3.4
155	METEOROLOGY & ATMOSPHERIC SCIENCES	10.9	0.5	0.047	61.3	10.5
156	ENGINEERING, AEROSPACE	2.5	0.2	0.095	68.4	2.2
157	MINERALOGY	6.9	0.4	0.060	61.4	6.6
158	MINING & MINERAL PROCESSING	4.0	0.3	0.069	65.5	3.7
159	GEOSCIENCES, MULTIDISCIPLINARY	7.3	0.4	0.055	62.7	6.9
	XV. AGRICULTURAL & ***ENVIRONMENT***					
160	AGRICULTURAL ENGINEERING	5.0	0.4	0.073	61.6	4.7
161	AGRICULTURE, MULTIDISCIPLINARY	6.8	0.3	0.045	63.8	6.6
162	AGRONOMY	5.8	0.3	0.050	62.9	5.5
163	LIMNOLOGY	9.7	0.8	0.078	60.8	9.3
164	SOIL SCIENCE	6.9	0.5	0.072	62.5	6.5
165	BIODIVERSITY CONSERVATION	8.8	0.4	0.046	62.1	8.5
166	ENVIRONMENTAL SCIENCES	8.9	0.5	0.056	60.1	8.8
167	ENVIRONMENTAL STUDIES	5.0	0.4	0.072	61.4	4.8
168	FOOD SCIENCE & TECHNOLOGY	7.1	0.5	0.075	61.9	6.7
169	NUTRITION & DIETETICS	11.4	0.4	0.037	61.3	11.1
170	AGRICULTURE, DAIRY & ANIMAL SCIENCE	5.4	0.3	0.051	66.5	4.9
171	HORTICULTURE	6.0	0.3	0.045	62.9	5.8
	XVI. BIOLOGY (ORGANISMIC AND SUPRAORGONISMIC LEVEL)					
172	ORNITHOLOGY	5.5	0.5	0.082	59.7	5.4
173	ZOOLOGY	7.5	0.5	0.068	61.8	7.1
174	ENTOMOLOGY	5.5	0.4	0.071	62.9	5.1
175	WATER RESOURCES	6.3	0.5	0.075	61.7	5.9
176	FISHERIES	7.1	0.8	0.115	59.3	6.9
177	MARINE & FRESHWATER BIOLOGY	8.2	0.9	0.115	59.2	7.9
178	MICROBIOLOGY	14.3	1.1	0.077	59.3	14.0
179	PARASITOLOGY	8.1	0.6	0.070	59.6	8.0
180	VIROLOGY	18.8	1.6	0.083	57.7	18.9
181	FORESTRY	7.2	0.6	0.089	60.0	7.0
182	MYCOLOGY	6.8	0.3	0.046	62.1	6.5
183	PLANT SCIENCES	9.6	0.3	0.029	60.1	9.8
184	ECOLOGY	11.4	1.0	0.087	59.7	11.0
185	VETERINARY SCIENCES	5.2	0.3	0.056	65.9	4.8
	XVII. MULTIDISCIPLINARY					
186	MULTIDISCIPLINARY SCIENCES	4.0	0.6	0.158	64.3	4.0

D. SOCIAL SCIENCES

	XVIII. SOCIAL SCIENCES, GENERAL					
187	CRIMINOLOGY & PENOLOGY	4.8	0.3	0.058	66.5	4.4
188	LAW	4.3	0.3	0.076	65.1	4.1
189	POLITICAL SCIENCE	3.3	0.4	0.119	65.5	3.2
190	PUBLIC ADMINISTRATION	3.6	0.3	0.075	66.2	3.3
191	ETHNIC STUDIES	2.5	0.3	0.115	65.7	2.4
192	FAMILY STUDIES	5.7	0.3	0.057	62.1	5.5
193	SOCIAL ISSUES	3.4	0.3	0.091	64.4	3.3
194	SOCIAL WORK	3.9	0.3	0.078	63.2	3.7
195	SOCIOLOGY	4.2	0.3	0.065	65.6	3.9
196	WOMEN'S STUDIES	4.1	0.2	0.061	63.8	3.8
197	EDUCATION & EDUCATIONAL RESEARCH	3.3	0.3	0.085	64.6	3.1
198	EDUCATION, SPECIAL	5.0	0.3	0.065	62.7	4.7
199	AREA STUDIES	1.9	0.3	0.157	67.0	1.8
200	GEOGRAPHY	5.8	0.3	0.057	60.5	5.7
201	PLANNING & DEVELOPMENT	4.4	0.3	0.059	61.3	4.4
202	TRANSPORTATION	5.3	0.4	0.079	61.8	5.0
203	URBAN STUDIES	4.4	0.3	0.068	61.7	4.2
204	ETHICS	3.3	0.3	0.092	65.6	3.0
205	MEDICAL ETHICS	5.2	0.4	0.075	62.1	4.9
206	ANTHROPOLOGY	4.4	0.3	0.074	66.3	4.1
207	COMMUNICATION	4.6	0.3	0.060	64.1	4.3
208	DEMOGRAPHY	5.5	0.3	0.053	61.8	5.3
209	HISTORY OF SOCIAL SCIENCES	2.1	0.3	0.140	69.2	1.8
210	INFORMATION SCIENCE & LIBRARY SCIENCE	4.1	0.4	0.103	65.2	3.9
211	INTERNATIONAL RELATIONS	2.9	0.4	0.134	65.4	2.8
212	LINGUISTICS	6.1	0.3	0.049	63.0	5.8
213	SOCIAL SCIENCES, INTERDISCIPLINARY	3.6	0.4	0.100	66.7	3.3
	XIX. ECONOMICS & BUSINESS					
214	AGRICULTURAL ECONOMICS & POLICY	3.8	0.3	0.082	63.9	3.5
215	ECONOMICS	4.6	0.3	0.074	61.9	4.6
216	INDUSTRIAL RELATIONS & LABOR	4.6	0.4	0.086	63.3	4.2
217	BUSINESS	6.7	0.3	0.047	64.0	6.4
218	BUSINESS, FINANCE	6.3	0.5	0.087	63.6	6.2
219	MANAGEMENT	6.4	0.4	0.055	63.5	6.2
	Mean			**0.071**	**62.2**	
	StDev			**0.043**	**3.0**	

Table 3B. Citation Inequality Decomposition at the Sub-field Level In the Fractional Case

	Quantiles	Within-group Term, *W*	Skew. of Sc. Term, *S*	*IDCP* Term	Total Citation Inequality	Percentages In %: (1)/(4)	(2)/(4)	(3)/(4)
		(1)	(2)	(3)	(4)	(5)	(6)	(7)
A. Raw Data	1,000	0.0030	0.7062	0.1552	0.8644	0.35	81.70	17.95
	[1, 660]			0.0463				5.36
	[661, 978]			0.0750				8.68
	[979, 1000]			0.0338				3.91

Table 4B. Exchange Rates, Standard Deviations, and Coefficients of Variation for the [661, 978] Interval. Multiplicative case.

		Exchange Rates	Standard Deviation	Coefficient of Variation	% of Citations	Exch. Rates Based on Mean Citations
		(1)	(2)	(3)	(4)	(5)
	A. LIFE SCIENCES					
	I. BIOSCIENCES					
1	BIOLOGY	10.5	0.4	0.035	63.8	10.0
2	BIOLOGY, MISCELLANEOUS	4.7	0.3	0.067	64.7	4.4
3	EVOLUTIONARY BIOLOGY	16.1	1.7	0.108	56.7	16.3
4	BIOCHEMICAL RESEARCH METHODS	11.5	0.6	0.054	54.7	12.4
5	BIOCHEMISTRY & MOLECULAR BIOLOGY	20.6	0.4	0.021	58.4	21.2
6	BIOPHYSICS	14	0.7	0.050	58.4	14.1
7	CELL BIOLOGY	27	1	0.038	60.4	27.5
8	GENETICS & HEREDITY	19.7	0.4	0.021	58.5	20.5
9	DEVELOPMENTAL BIOLOGY	24.4	0.5	0.021	60.4	24.6
	II. BIOMEDICAL RESEARCH					
10	PATHOLOGY	11.7	0.3	0.024	62	11.5
11	ANATOMY & MORPHOLOGY	7.8	0.5	0.064	60.9	7.6
12	ENGINEERING, BIOMEDICAL	9.6	0.5	0.048	61.2	9.2
13	BIOTECHNOLOGY & APPLIED MICROBIOLOGY	11.6	0.3	0.022	57.9	12.1
14	MEDICAL LABORATORY TECHNOLOGY	8.1	0.3	0.031	61.3	8.0
15	MICROSCOPY	8.5	0.6	0.068	60.6	8.3
16	PHARMACOLOGY & PHARMACY	10.7	0.4	0.041	59.8	10.6
17	TOXICOLOGY	9.6	0.6	0.067	59.2	9.5
18	PHYSIOLOGY	14.1	1.4	0.101	59.3	13.7
19	MEDICINE, RESEARCH & EXPERIMENTAL	15.7	2.8	0.180	59.9	17.2
	III. CLINICAL MEDICINE I (INTERNAL)					
20	CARDIAC & CARDIOVASCULAR SYSTEMS	14.9	1.1	0.076	61.3	15.2
21	RESPIRATORY SYSTEM	13.5	0.6	0.042	60.6	13.2
22	ENDOCRINOLOGY & METABOLISM	16.7	1.1	0.066	58.2	16.9
23	ANESTHESIOLOGY	9.4	0.3	0.032	62.8	8.9
24	CRITICAL CARE MEDICINE	14.6	0.4	0.030	61.5	14.2
25	EMERGENCY MEDICINE	5.8	0.3	0.050	62.2	5.6
26	GASTROENTEROLOGY & HEPATOLOGY	13.7	0.4	0.027	60.4	13.8
27	MEDICINE, GENERAL & INTERNAL	12.1	5	0.411	52.2	16.9
28	TROPICAL MEDICINE	7.2	0.5	0.069	62.1	6.8
29	HEMATOLOGY	21.9	0.4	0.020	61	21.8
30	ONCOLOGY	18	0.5	0.027	58.8	18.3
31	ALLERGY	12.2	0.4	0.033	62.7	11.7
32	IMMUNOLOGY	17.8	0.3	0.016	58.9	18.3
33	INFECTIOUS DISEASES	15.3	0.9	0.060	59.4	15.2

	IV. CLINICAL MEDICINE II (NON-INTERNAL)					
34	GERIATRICS & GERONTOLOGY	11.1	0.6	0.054	61.5	10.7
35	OBSTETRICS & GYNECOLOGY	9.2	0.4	0.042	62.1	8.8
36	ANDROLOGY	7.4	0.6	0.079	60.1	7.2
37	REPRODUCTIVE BIOLOGY	12.6	1.1	0.088	58.7	12.4
38	GERONTOLOGY	10	0.4	0.038	63.3	9.4
39	DENTISTRY & ORAL SURGERY	7.2	0.5	0.073	60.6	7.0
40	DERMATOLOGY	8.1	0.3	0.036	62.1	7.8
41	UROLOGY & NEPHROLOGY	12.4	0.3	0.022	61.9	12.0
42	OTORHINOLARYNGOLOGY	6.1	0.4	0.069	62.4	5.7
43	OPHTHALMOLOGY	9.5	0.3	0.030	61.3	9.3
44	INTEGRATIVE & COMPLEMENTARY MEDICINE	6.2	0.6	0.090	61.2	5.9
45	CLINICAL NEUROLOGY	12.4	0.3	0.021	61.4	12.2
46	PSYCHIATRY	13.1	0.3	0.020	62	12.8
47	RADIOLOGY, NUCLEAR MED. & MED. IMAGING	10.4	0.3	0.025	61.4	10.3
48	ORTHOPEDICS	7.9	0.3	0.038	61.4	7.7
49	RHEUMATOLOGY	14.6	0.6	0.038	59.7	14.6
50	SPORT SCIENCES	8.2	0.5	0.056	62.5	7.7
51	SURGERY	8.6	0.2	0.028	62	8.4
52	TRANSPLANTATION	9.3	0.3	0.029	61.9	9.1
53	PERIPHERAL VASCULAR DISEASE	20.4	0.3	0.013	60.3	20.5
54	PEDIATRICS	7.7	0.3	0.035	61.8	7.5
	V. CLINICAL MEDICINE III					
55	HEALTH CARE SCIENCES & SERVICES	7.8	0.4	0.049	60.7	7.6
56	HEALTH POLICY & SERVICES	8.2	0.3	0.039	59.3	8.2
57	MEDICINE, LEGAL	5.8	0.4	0.069	60.5	5.6
58	NURSING	4.4	0.4	0.091	62.4	4.1
59	PUBLIC, ENV. & OCCUPATIONAL HEALTH	9.6	0.3	0.035	60.7	9.5
60	REHABILITATION	5.9	0.4	0.060	62.5	5.6
61	SUBSTANCE ABUSE	10	0.9	0.090	59.1	9.8
62	EDUCATION, SCIENTIFIC DISCIPLINES	4	0.3	0.071	64.8	3.8
63	MEDICAL INFORMATICS	5.7	0.3	0.046	61.6	5.6
	VI. NEUROSCIENCES & BEHAVIORAL					
64	NEUROIMAGING	14.6	0.4	0.026	63.1	14.0
65	NEUROSCIENCES	17	0.5	0.029	59.5	17.1
66	BEHAVIORAL SCIENCES	11.5	1.3	0.115	56	11.7
67	PSYCHOLOGY, BIOLOGICAL	9.9	0.8	0.084	57.3	10.0
68	PSYCHOLOGY	10.6	0.7	0.069	60.1	10.3
69	PSYCHOLOGY, APPLIED	6.5	0.4	0.063	61.9	6.2
70	PSYCHOLOGY, CLINICAL	10	0.4	0.038	61.2	9.8
71	PSYCHOLOGY, DEVELOPMENTAL	10.4	0.5	0.052	60.8	10.1
72	PSYCHOLOGY, EDUCATIONAL	7.1	0.3	0.043	64	6.7
73	PSYCHOLOGY, EXPERIMENTAL	10.2	0.4	0.042	61	10.0
74	PSYCHOLOGY, MATHEMATICAL	7	0.3	0.038	61	6.9
75	PSYCHOLOGY, MULTIDISCIPLINARY	6.4	0.6	0.092	62.6	6.4
76	PSYCHOLOGY, PSYCHOANALYSIS	3.8	0.4	0.100	66.3	3.5
77	PSYCHOLOGY, SOCIAL	8.3	0.3	0.031	61.6	8.1

78	SOCIAL SCIENCES, BIOMEDICAL	7.4	0.3	0.039	60.7	7.3
	B. PHYSICAL SCIENCES					
	VII. CHEMISTRY					
79	CHEMISTRY, MULTIDISCIPLINARY	12	1.3	0.108	65	11.7
80	CHEMISTRY, INORGANIC & NUCLEAR	9.1	0.6	0.062	61.6	8.7
81	CHEMISTRY, ANALYTICAL	10	0.5	0.046	60.6	9.8
82	CHEMISTRY, APPLIED	7.7	0.5	0.063	61.9	7.3
83	ENGINEERING, CHEMICAL	6	0.3	0.045	63.9	5.7
84	CHEMISTRY, MEDICINAL	9.8	0.8	0.078	59	9.7
85	CHEMISTRY, ORGANIC	10.7	1	0.090	59.1	10.5
86	CHEMISTRY, PHYSICAL	10.5	0.4	0.043	60	10.4
87	ELECTROCHEMISTRY	10.4	0.7	0.072	60.6	10.0
88	POLYMER SCIENCE	8.3	0.3	0.031	61.3	8.1
	VIII. PHYSICS					
89	PHYSICS, MULTIDISCIPLINARY	10.1	1.7	0.169	62.2	10.6
90	SPECTROSCOPY	7.7	0.3	0.043	61.8	7.4
91	ACOUSTICS	5.6	0.3	0.052	62.7	5.3
92	OPTICS	7.3	0.3	0.038	62.8	7.1
93	PHYSICS, APPLIED	7.5	0.4	0.049	60.9	7.6
94	PHYSICS, ATOMIC, MOLECULAR & CHEMICAL	11.1	0.8	0.071	59.1	11.0
95	THERMODYNAMICS	4.8	0.4	0.081	61.7	4.6
96	PHYSICS, MATHEMATICAL	7.5	0.3	0.037	61.6	7.4
97	PHYSICS, NUCLEAR	6.6	0.4	0.067	63.3	6.4
98	PHYSICS, PARTICLES & SUB-FIELDS	11.1	1.2	0.106	60.7	11.6
99	PHYSICS, CONDENSED MATTER	7.5	0.3	0.039	62	7.4
100	PHYSICS OF SOLIDS, FLUIDS & PLASMAS	9.4	0.6	0.064	60	9.2
101	CRYSTALLOGRAPHY	5.2	0.2	0.046	56.4	5.6
	IX. SPACE SCIENCES					
102	ASTRONOMY & ASTROPHYSICS	14.9	0.3	0.018	60.7	14.9
	X. MATHEMATICS					
103	MATHEMATICS, APPLIED	3.7	0.3	0.075	65	3.5
104	STATISTICS & PROBABILITY	5.4	0.5	0.097	54.1	6.2
105	MATH., INTERDISCIPLINARY APPLICATIONS	5.6	0.2	0.044	61.6	5.5
106	SOCIAL SCIENCES, MATHEMATICAL METHODS	5.6	0.3	0.047	61.4	5.5
107	PURE MATHEMATICS	2.8	0.2	0.087	66	2.6
	XI. COMPUTER SCIENCE					
108	COMP. SCIENCE, ARTIFICIAL INTELLIGENCE	4.8	0.5	0.107	63.4	4.8
109	COMPUTER SCIENCE, CYBERNETICS	3.7	0.4	0.102	67.1	3.4
110	COMP. SCIENCE, HARDWARE & ARCHITECTURE	3.9	0.5	0.123	62.9	4.0
111	COMPUTER SCIENCE, INFORMATION SYSTEMS	4.3	0.7	0.154	62.5	4.5
112	COMP. SC., INTERDISCIPLINARY APPLICATIONS	5.7	0.6	0.099	56.6	6.3
113	COMP. SCIENCE, SOFTWARE ENGINEERING	3.7	0.4	0.114	65	3.5
114	COMPUTER SCIENCE, THEORY & METHODS	2.9	0.4	0.130	65.6	2.8

115	MATHEMATICAL & COMPUTATIONAL BIOLOGY	9.8	0.5	0.047	49.7	12.2
	C. OTHER NATURAL SCIENCES					
	XII. ENGINEERING					
116	ENGINEERING, ELECTRICAL & ELECTRONIC	4.8	0.4	0.077	63	4.7
117	TELECOMMUNICATIONS	3.7	0.5	0.147	63.6	3.8
118	CONSTRUCTION & BUILDING TECHNOLOGY	3.5	0.3	0.088	65.5	3.2
119	ENGINEERING, CIVIL	3.4	0.3	0.087	66.3	3.2
120	ENGINEERING, ENVIRONMENTAL	9	0.3	0.034	62.5	8.7
121	ENGINEERING, MARINE	1.5	0.3	0.210	71.5	1.4
122	TRANSPORTATION SCIENCE & TECHNOLOGY	2.1	0.5	0.233	70.9	1.9
123	ENGINEERING, INDUSTRIAL	3.3	0.3	0.088	66.2	3.0
124	ENGINEERING, MANUFACTURING	3.6	0.3	0.087	65.3	3.2
125	ENGINEERING, MECHANICAL	4	0.2	0.060	63.9	3.8
126	MECHANICS	5.2	0.3	0.049	63.4	4.9
127	ROBOTICS	3.7	0.3	0.069	65	3.5
128	INSTRUMENTS & INSTRUMENTATION	5.2	0.2	0.046	64.4	4.9
129	IMAGING SCIENCE & PHOTOGR. TECHNOLOGY	7.5	0.4	0.058	63.8	7.2
130	ENERGY & FUELS	5.2	0.3	0.056	64.5	4.9
131	NUCLEAR SCIENCE & TECHNOLOGY	4.4	0.3	0.059	62.9	4.2
132	ENGINEERING, PETROLEUM	1.7	0.4	0.257	73.5	1.5
133	AUTOMATION & CONTROL SYSTEMS	4.1	0.2	0.060	64.5	3.8
134	ENGINEERING, MULTIDISCIPLINARY	3.9	0.4	0.101	65.9	3.6
135	ERGONOMICS	4.8	0.4	0.080	62.4	4.5
136	OPERATIONS RES. & MANAGEMENT SCIENCE	4	0.2	0.061	63.9	3.8
	XIII. MATERIALS SCIENCE					
137	MATERIALS SCIENCE, MULTIDISCIPLINARY	6.5	0.4	0.061	60.6	6.6
138	MATERIALS SCIENCE, BIOMATERIALS	13	1.1	0.084	59.1	12.8
139	MATERIALS SCIENCE, CERAMICS	4.8	0.4	0.075	68.1	4.3
140	MAT. SC., CHARACTERIZATION & TESTING	2.2	0.4	0.189	69.5	2.0
141	MATERIALS SCIENCE, COATINGS & FILMS	7.5	0.5	0.065	61.4	7.2
142	MATERIALS SCIENCE, COMPOSITES	3.5	0.3	0.084	65.1	3.3
143	MATERIALS SCIENCE, PAPER & WOOD	3	0.3	0.091	68	2.6
144	MATERIALS SCIENCE, TEXTILES	2.9	0.3	0.089	65.5	2.7
145	METALL. & METALLURGICAL ENGINEERING	4.6	0.4	0.082	64.5	4.4
146	NANOSCIENCE & NANOTECHNOLOGY	8.2	0.4	0.044	59.6	8.4
	XIV. GEOSCIENCES					
147	GEOCHEMISTRY & GEOPHYSICS	9.8	0.6	0.060	61.7	9.4
148	GEOGRAPHY, PHYSICAL	9	0.8	0.088	59.9	8.7
149	GEOLOGY	8	0.4	0.055	62.7	7.6
150	ENGINEERING, GEOLOGICAL	3.7	0.3	0.088	62.5	3.5
151	PALEONTOLOGY	6.4	0.4	0.055	63.1	6.0
152	REMOTE SENSING	7.4	0.3	0.043	60.6	7.3
153	OCEANOGRAPHY	10	0.9	0.090	61.2	9.5
154	ENGINEERING, OCEAN	3.8	0.4	0.098	64.8	3.6
155	METEOROLOGY & ATMOSPHERIC SCIENCES	10.6	0.4	0.037	61.3	10.3

156	ENGINEERING, AEROSPACE	2.6	0.2	0.091	68.7	2.3
157	MINERALOGY	7.2	0.4	0.060	61.7	6.8
158	MINING & MINERAL PROCESSING	4.1	0.3	0.065	65.8	3.9
159	GEOSCIENCES, MULTIDISCIPLINARY	7.3	0.4	0.050	62.6	6.9
	XV. AGRICULTURAL & ENVIRONMENT					
160	AGRICULTURAL ENGINEERING	4.9	0.4	0.072	62	4.7
161	AGRICULTURE, MULTIDISCIPLINARY	6.9	0.3	0.038	64.7	6.4
162	AGRONOMY	5.9	0.3	0.046	63	5.6
163	LIMNOLOGY	9.5	0.6	0.065	61	9.2
164	SOIL SCIENCE	6.9	0.5	0.074	62.1	6.5
165	BIODIVERSITY CONSERVATION	8.8	0.3	0.037	62.7	8.4
166	ENVIRONMENTAL SCIENCES	8.9	0.5	0.051	60.8	8.7
167	ENVIRONMENTAL STUDIES	4.9	0.3	0.071	61.7	4.7
168	FOOD SCIENCE & TECHNOLOGY	7.1	0.5	0.067	61.8	6.8
169	NUTRITION & DIETETICS	11.4	0.3	0.030	61.3	11.1
170	AGRICULTURE, DAIRY & ANIMAL SCIENCE	5.4	0.3	0.048	65.9	5.0
171	HORTICULTURE	6.2	0.3	0.044	62.9	6.0
	XVI. BIOLOGY (ORGANISMIC AND SUPRAORGONISMIC LEVEL)					
172	ORNITHOLOGY	5.5	0.4	0.077	59.8	5.4
173	ZOOLOGY	7.5	0.5	0.065	61.4	7.2
174	ENTOMOLOGY	5.5	0.4	0.067	63	5.1
175	WATER RESOURCES	6.2	0.4	0.068	62.2	5.8
176	FISHERIES	7.1	0.8	0.110	60	6.8
177	MARINE & FRESHWATER BIOLOGY	8.2	0.9	0.113	59.2	7.9
178	MICROBIOLOGY	14.3	1	0.071	58.9	14.2
179	PARASITOLOGY	8.1	0.6	0.072	60	7.9
180	VIROLOGY	18.7	1.5	0.082	57.6	18.8
181	FORESTRY	7	0.6	0.079	60.2	6.8
182	MYCOLOGY	6.8	0.3	0.046	62.3	6.5
183	PLANT SCIENCES	9.6	0.3	0.027	60.7	9.6
184	ECOLOGY	11.4	1	0.085	59.7	11.1
185	VETERINARY SCIENCES	5.2	0.3	0.054	65.4	4.8
	XVII. MULTIDISCIPLINARY					
186	MULTIDISCIPLINARY SCIENCES	4.1	0.6	0.161	64.2	4.1
	D. SOCIAL SCIENCES					
	XVIII. SOCIAL SCIENCES, GENERAL					
187	CRIMINOLOGY & PENOLOGY	4.9	0.3	0.065	66.5	4.5
188	LAW	4.4	0.4	0.083	64.7	4.2
189	POLITICAL SCIENCE	3.3	0.4	0.119	65.7	3.2
190	PUBLIC ADMINISTRATION	3.7	0.3	0.075	65.9	3.4
191	ETHNIC STUDIES	2.6	0.3	0.103	66	2.4
192	FAMILY STUDIES	5.8	0.3	0.055	62	5.6
193	SOCIAL ISSUES	3.6	0.3	0.088	65.5	3.4
194	SOCIAL WORK	3.9	0.3	0.069	63.4	3.6
195	SOCIOLOGY	4.2	0.3	0.067	65.1	4.0
196	WOMEN'S STUDIES	4	0.3	0.063	64	3.8

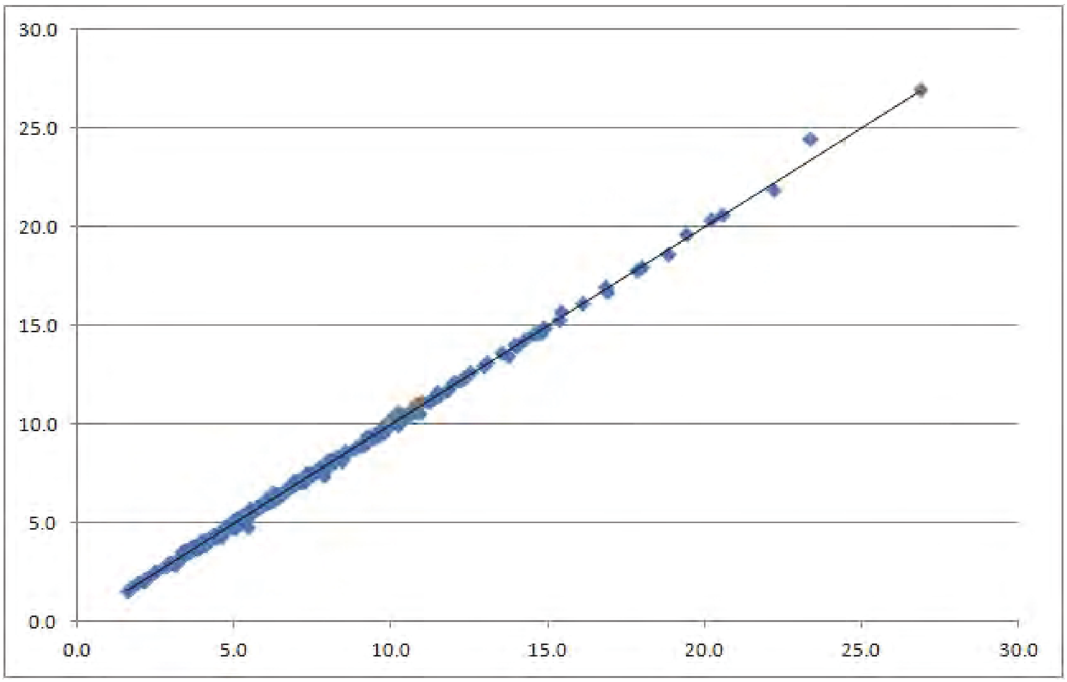
197	EDUCATION & EDUCATIONAL RESEARCH	3.3	0.3	0.088	64.3	3.1
198	EDUCATION, SPECIAL	5.1	0.3	0.059	62.5	4.9
199	AREA STUDIES	2	0.3	0.154	67.4	1.8
200	GEOGRAPHY	5.8	0.3	0.054	60.8	5.7
201	PLANNING & DEVELOPMENT	4.3	0.3	0.060	62.4	4.2
202	TRANSPORTATION	5.1	0.4	0.073	62.2	4.9
203	URBAN STUDIES	4.3	0.3	0.064	62.3	4.1
204	ETHICS	3.5	0.3	0.080	65.3	3.2
205	MEDICAL ETHICS	5.2	0.4	0.071	62.1	4.9
206	ANTHROPOLOGY	4.3	0.3	0.075	65.9	4.0
207	COMMUNICATION	4.3	0.3	0.065	63.4	4.0
208	DEMOGRAPHY	5.6	0.3	0.048	61.3	5.5
209	HISTORY OF SOCIAL SCIENCES	2.1	0.3	0.145	69.1	1.8
210	INFORMATION SCIENCE & LIBRARY SCIENCE	3.9	0.5	0.127	64.1	3.8
211	INTERNATIONAL RELATIONS	2.9	0.4	0.140	65.5	2.9
212	LINGUISTICS	6	0.3	0.046	63.5	5.7
213	SOCIAL SCIENCES, INTERDISCIPLINARY	3.5	0.3	0.098	66.1	3.3
	XIX. ECONOMICS & BUSINESS					
214	AGRICULTURAL ECONOMICS & POLICY	3.8	0.3	0.073	63.6	3.6
215	ECONOMICS	4.6	0.4	0.077	62	4.6
216	INDUSTRIAL RELATIONS & LABOR	4.5	0.3	0.077	64.1	4.1
217	BUSINESS	6.7	0.4	0.056	64.3	6.4
218	BUSINESS, FINANCE	6.4	0.6	0.094	64.3	6.3
219	MANAGEMENT	6.4	0.4	0.061	63.6	6.2
	Mean			**0.07**	**62.2**	

Table 5B. Citation Inequality Decomposition Sat the Sub-field level. The Multiplicative Case.

	Quantiles	Within-group Term, $\mathbb{W}$	Skew. of Sc. Term, $\mathbb{S}$	$IDCP$ Term	Total Citation Inequality	Percentages In %: (1)/(4)	(2)/(4)	(3)/(4)
		(1)	(2)	(3)	(4)	(5)	(6)	(7)
A. Raw Data	All quantiles	0.0030	0.6950	0.1544	0.8524	0.35	81.54	18.11
	[1, 660]			0.0469				5.50
	[661, 978]			0.0766				8.98
	[979, 1000]			0.0310				3.63
B. Sub-field ER Normalization	All quantiles	0.0030	0.7212	0.0268	0.7510	0.41	96.03	3.57
	[1, 660]			0.0160				2.13
	[661, 978]			0.0023				0.31
	[979, 1000]			0.0085				1.13
C. Sub-field Mean Normalization	All quantiles	0.0029	0.7168	0.0243	0.7440	0.39	96.34	3.27
	[1, 660]			0.0164				2.20
	[661, 978]			0.0023				0.31
	[979, 1000]			0.0056				0.76

Figure 4B. A Comparison at the Sub-field level of Exchange Rates in the Fractional *versus* the Multiplicative Case

Exchange Rates. Multiplicative Case



Exchange Rates. Fractional Case

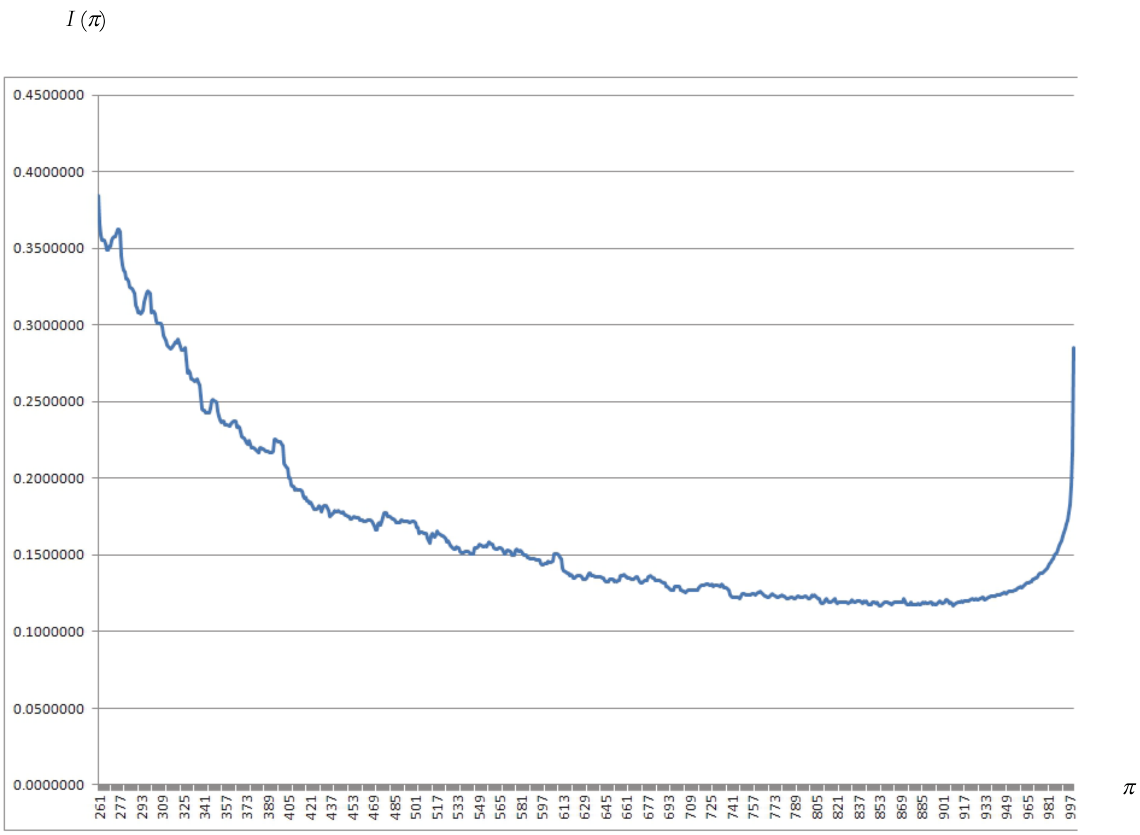


Figure 1B. Citation Inequality Due to Differences in Citation Practices, $I(\pi)$ versus π. Raw Data

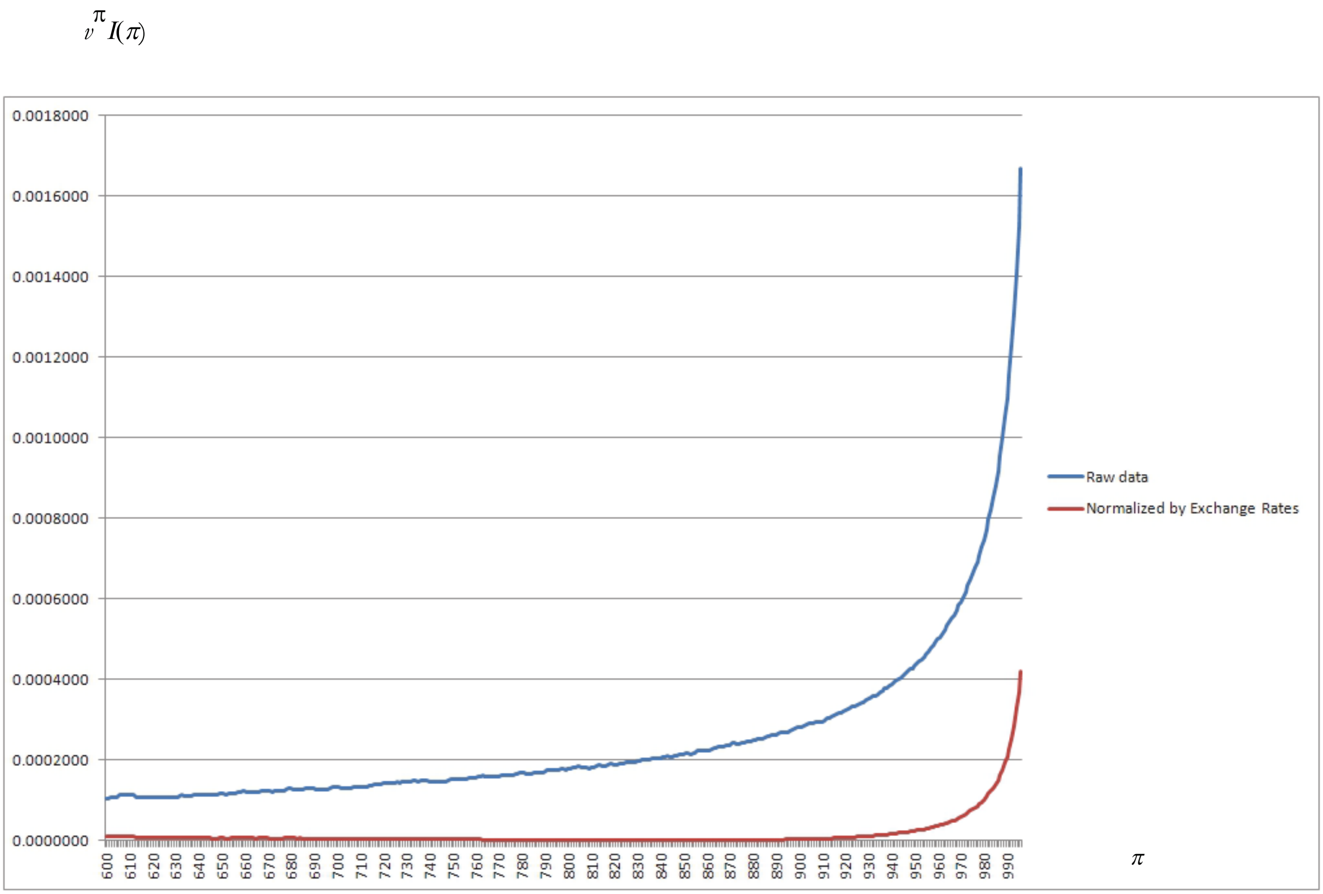
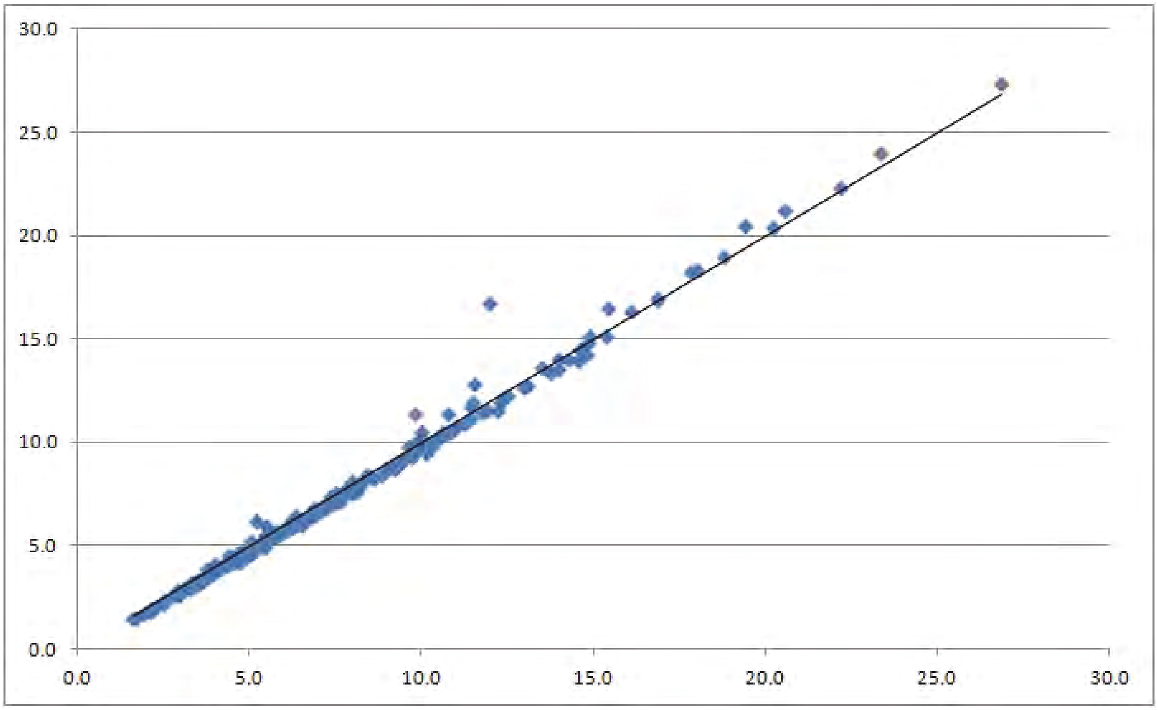


Figure 2B. Weighted Citation Inequality Due to Differences in Citation Practices, v^{π} $I(\pi)$ vs. π. Raw *vs.* Normalized Data

Figure 3B. A Comparison at the Sub-field Level of the Estimated ERs Over the [661, 978] Interval *versus* the Exchange Rates Based on Mean Citations. The Fractional Case.

Exchange Rates Based on Mean Citations



Estimated Exchange Rates