ESTUDIOS DE ECONOMIA

VOLUMEN 52 • Nº2 / DICIEMBRE 2025

ARTÍCULOS

UNIVERSIDAD DE CHILE
FACULTAD DE ECONOMIA Y NEGOCIOS
DEPARTAMENTO DE ECONOMIA

ESTUDIOS DE ECONOMIA

POLÍTICA EDITORIAL

La Revista ESTUDIOS DE ECONOMÍA es publicada dos veces al año por el Departamento de Economía de la Universidad de Chile. Su propósito es presentar artículos acerca de investigaciones empíricas y ensayos teóricos que sean de interés general para economistas.

Todo trabajo será evaluado para su publicación por dos comentaristas anónimos.

Como parte de nuestra política editorial, nos comprometemos a enviar una respuesta acerca de la situación de cada trabajo en un plazo no superior a los dos meses desde su recepción. De no hacerlo, los autores se encuentran en plena libertad de considerar otra revista para la publicación de su artículo.

Toda correspondencia debe dirigirse a:

EDITORIAL POLICY

ESTUDIOS DE ECONOMIA is published two times a year by the Department of Economics at the University of Chile. Its purpose is to publish empirical as well as theoretical papers in every field of Economics.

All papers must go through refereeing process by two anonimous referees.

As part of our editorial policy, an answer regarding the status of a paper will not exceed two months from its reception. If that were not the case, the authors are free to send their paper for consideration for publication elsewhere.

All correspondence should be sent to:

Revista, *Estudios de Economía*
http://www.estudiosdeeconomia.uchile.cl
Publicación Nº 263 del
Departamento de Economía, Universidad de Chile
Diagonal Paraguay 257
Registro Propiedad Intelectual Nº 112.761
Santiago - CHILE
ede@econ.uchile.cl

Estudios de Economía is indexed and abstracted in:

- Social Sciences Citation Index.
- Social Scisearch.
- Journal Citation Reports/Social Sciences Edition.

UNIVERSIDAD DE CHILE • FACULTAD DE ECONOMIA Y NEGOCIOS

ESTUDIOS DE ECONOMÍA

VOLUMEN 52 - Nº2 ISSN 0304-2758 DICIEMBRE 2025

ARTÍCULOS

Departamento de Economía

On the Efficiency and Stability of a Two-Way Flow Network with Small Decay: A Note*

Sobre la eficiencia y la estabilidad en una red de flujo bidireccional con pequeño decaimiento: una nota

BANCHONGSAN CHAROENSOOK**

Abstract

The seminal two-way flow strategic network formation model of Bala and Goyal (2000b) exhibits a substantial tension between stability and efficiency. In this note, I show that despite this tension every link receiver in a Nash network serves as an efficient transmitter of information assuming a small degree of information decay. Thus, a strategic decision of every link sender, who always forms links with efficient link receivers for his own interest, does in part lead to a socially desirable outcome. I also show how this finding can potentially refine some results in the literature.

Key words: *Network Formation, Nash Network, Two-way Flow Network, Agent Heterogeneity, Efficient Network.*

JEL Classification: *C72, D85.*

* At Keimyung Adams College, I thank my former assistants - Youngju Cho, Sua Lee, Jaetong Son, Hyeonmin Kim, Yudam Jeong and Hayeong Lee - for their excellent help that substantially saved my time during this research. I also wish to thank two anonymous reviewers for their helpful comments. This research is supported by Keimyung University Outstanding Research Grant Number 20230335. ORCID: https://orcid.org/0000-0001-7225-8455

** Faculty of Business and Management, American University of Phnom Penh, Phnom Penh, Cambodia. Email: b.charoensook@aupp.edu.kh.

DOI: http://dx.doi.org/10.4067/S0718-52862025000200235

Received: August, 2023 Accepted: January, 2025

Resumen

El modelo seminal de formación de redes estratégicas de flujo bidireccional de Bala y Goyal (2000b) presenta una tensión considerable entre la estabilidad y la eficiencia. En esta nota, muestro que, a pesar de dicha tensión, cada receptor de enlace en una red de Nash actúa como un transmisor eficiente de información si se asume un pequeño grado de decaimiento de la información. De este modo, la decisión estratégica de cada emisor de enlace, quien siempre forma enlaces con receptores de enlace eficientes en su propio beneficio, conduce en parte a un resultado socialmente deseable. Además, muestro cómo este hallazgo puede refinar algunos resultados existentes en la literatura.

Palabras clave: *Formación de redes, red de Nash, Red bidireccional, Agentes heterogéneos, Red eficiente.*

Clasificación JEL: *C72, D85.*

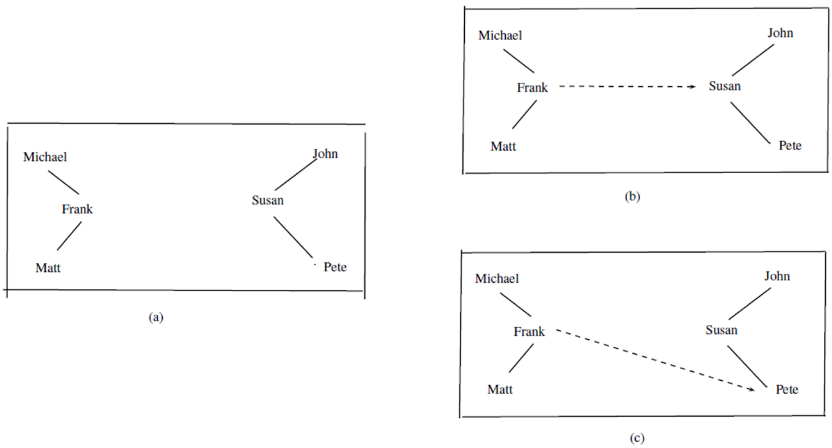
1. Introduction

In a social network, information transmission tends to be imperfect. In the literature of strategic network formation, a rigorously studied form of imperfect information transmission is the so-called 'small decay'. It captures the idea that the worth of information decays as it traverses through each link. Yet this decay is sufficiently small that, if a chain between two agents exists, then neither agent finds an incentive to shorten the distance through a costly link formation. This assumption of 'small decay' is rigorously studied in the context of two-way flow model of network formation with nonrival information by De Jaegher and Kamphorst (2015), which is an extended model of network formation proposed in the seminal paper of Bala and Goyal (2000b). Their major novel findings, which allow them to finely characterize the equilibrium networks, are; (i) best-informed agents, defined as agents who received more information than others, are attractive as link receivers and (ii) these best-informed agents are located 'in the middle' of other agents.

In this note, I complement these novel findings by showing that every best-informed agent, in addition to being an optimal choice as a link receiver from a strategic perspective of self-interest agents, is also an agent that transmits information *most efficiently* in the network. That is, within a group of agents M if a link sender i's best response is to choose a best-informed agent $j \in M$ as a link receiver, then j is also the agent that maximizes the total information that flows within the network. These results are established as Proposition 1 and Remark 2 in this note. See Figure Figure 1 in the Appendix for an intuitive, informal example of this insight[1].

FIGURE 1

EXAMPLE 1



Notes: On the left (picture (a)), two groups of agents are disconnected. If the decay is small, then within the group of Susan, Pete and John it can be shown that Susan, who is in the middle, possesses more informational quantity than Pete and John do (see lemma 1 in de jaegher and kamphorst (2015)). Consequently, on one hand, if Frank wants to acquire information with the group of Susan, John and Pete, then for the benefit of his own interest Frank will choose Susan as his partner, as shown on the upper right (picture (b)), rather than John or Pete, as shown on the lower right (picture (c)). On the other hand, proposition 1 in this note further shows that Susan - rather than John and Pete - is also an agent that efficiently transmits information in the network.

I provide a brief literature review and threefold contributions of this note as follows. First, existing literature in the strand of two-way flow network formation shows that there is a substantial tension between efficient network and equilibrium networks. Specifically, within the context of the aforementioned two-way flow model of network formation with nonrival information Proposition 1 and 2 in De Jaegher and Kamphorst (2015) show that the set of Nash networks can be large and does contain networks that have long diameters,

[1] This Figure is inspired by Figure 1 in De Jaegher and Kamphorst (2015).

networks can be large and does contain networks that have long diameters, while Proposition 5.5 in Bala and Goyal (2000b) shows that a star - a network with diameter of only 2 - is a unique minimal nonempty efficient network. Thus, considering the differences in terms of network architectures and diameters the tension between efficient networks and Nash networks is quite extremE[2 3]. However, despite this tension between stability and efficiency to my knowledge this note is the first work in the literature that provides a connection between efficient network and Nash network by precisely pinpoints the source of inefficiency in a Nash network; identities of link senders.

Second, my result shows that in a Nash network a strategic choice made by self-interested agents who act as link senders, which results in every link receiver being efficient, does in part lead to a socially desirable outcome. Put differently, the fact that each link sender has to bear the entire link formation cost does not cause a link sender to choose an inefficient link receiver. Thus, by assuming a less general form of payoff function my result stands in contrast to the remark of Jackson (2008), "there are still inefficiencies, most notably since only one player bears the cost of a link while many players benefit from its existence."

Third, the above result also becomes a technical tool that refines some existing results in the literature. By incorporating small decay assumption, multiple efficient networks (assuming no decay) can be ranked in terms of efficiency of information flow. Consequently, a large set of efficient networks - which is a technical difficulty in some existing work - can be substantially narrowed down. At the same time, the incorporation of small decay assumption helps understand which efficient networks under the assumption of perfect communication are likely to be resilient to a small degree of imperfect communication. In Section 4, I show how to use this technical tool to refine the results of Unlu (2018), which finds that agent heterogeneity can lead to a large set of efficient

---

2 A similar tension also exists if no decay and agent homogeneity or agent heterogeneity as in Unlu (2018) is assumed. For more details, Unlu (2018) elaborates on this conflict found in Jackson and Wolinsky (1996), Bala and Goyal (2000b) and Bala and Goyal (2000a).

3 Indeed, such a tension has also been mentioned by Breitmoser and Vorjohann (2013). In par ticular, Breitmoser and Vorjohann (2013) remark that stars or complete networks are efficient across various models of network formation, including two-way flow model with bilateral link consent by Jackson and Wolinsky (1996), two-way flow model with cut-off decay (Hojman and Szeidl (2008)), model with far-sighted players (Dutta et al. (2005)), model with endogenous link strength (Bloch and Dutta (2009)) and model with transfer payments between players (Bloch and Jackson (2007)). Breitmoser and Vorjohann (2013) complement the literature by showing that substantially differ ent architectures of networks- redundant, incomplete and circular networks- are efficient if noisy communication is assumed. See the first two paragraphs in Breitmoser and Vorjohann (2013) for a comprehensive literature review.

networks that contains long diameter networks.

This note proceeds as follows. Section 2 introduces the model and two important concepts - efficient link receiver/transmitter and best-informed agent. Section 3 provides main results. Section 4 provides applications of the main results to the literature. Section 5 concludes with remarks on further potential studies.

2. THE MODEL

This note primarily follows the notation of De Jaegher and Kamphorst (2015), since it is the paper that this note seeks to complement.

2.1 Link Establishment And Individual's Strategy

Let $N = \{1,...,n\}$ be the set of all agents. An agent $i \in N$ can form a link with another agent j without j's consent. ij denotes such a link. The set of all possible links that i forms is $L_i = \{ij; j \in N \setminus \{i\}\}$. The set of all possible links is $L \equiv \cup_{i \in N} L_i$. $g_i \subset L_i$ is a *strategy* of i and $g = \cup_{i \in N} g_i$ is a *strategy profile*. A strategy space G is $G \equiv 2^L$. Pictorially, a strategy profile g is also a network, where an arrow from agent i to j indicates that $ij \in g_i$.

2.2 Information Flow

Information flow is two-way in the sense that it flows between two agents regardless of who sponsors the link, hence the term 'two-way flow model'. Let $\overline{ij} \in g$ represents that either $ij \in g$ or $ji \in g$. Information can also flow via a *chain*. A chain between i and j in a network g, denoted by $P_{ij}(g)$, is a sequence of agents $\{\overline{i_0 i_1},...,\overline{i_{k-1} i_k}\}$ such that $i_0 = i, i_k = j$. If there is a chain between i and j, we say that i and j are connected. A *shortest chain* between i and j is, of course, the chain(s) between i and j with the least amount of links. The distance between i and j, denoted by $d_{ij}(g)$ is defined as the amount of links of the shortest chain(s). If $j = i$ then we assume, following the literature, that the distance between i and himself is 0. If j and i are not connected, then we set $d_{ij}(g) = \infty$.

2.3 Information Decay

Let $\sigma \in [0,1]$ denote the *decay factor*. For example, if the value of information that an agent j possesses is 1 and the distance between i and j is k then the information that i receives from j in g is σ^k.

2.4 Small Decay Assumption

Suppose that information of j flows to i via a multi-link chain, then i can improve the information flow by establishing a link that results in a shorter chain. Such an incentive arises if the improvement in terms of information flow exceeds the increasing link establishment cost. However, if the decay factor σ is close to 1 then the improvement in terms of information flow becomes marginal and, consequently, such an incentive to establish a link diminishes. By the same analogy, from an efficiency perspective the benefits to all agents in the network relative to the cost of establishing an extra chain also diminishes if the decay is sufficiently small. As a result, there is at most only one chain between any pair of agents. This small decay assumption is assumed in De Jaegher and Kamphorst (2015) and will be assumed throughout this note. See Lemma 1 that precedes Proposition 1 in the next section.

2.5 Network-Related Notations

A subnetwork of g is a network g' such that $g' \subset g$. A network is said to be *connected* if there is a chain between every pair of agents in the network. g', a subnetwork of g, is said to be a component of g if g' is a maximal connected subnetwork of g. A network is *empty* if no agent forms a link. A non-empty component of a network or a network is *minimal* if there is at most one chain between any pair of agents in the network. An agent i is called a *link sender* (*receiver*) if there is a link $xy \in g$ such that $x = i$ ($y = i$). A minimally connected network is a *star* if there is an agent i such that $ij \in g$ for every $j \neq i >$ but $jk \notin g$ for every $j,k \neq i$.

Next, I introduce some notations concerning information flow. Let g be a minimally connected network. Due to the fact that there is only one chain between every pair of agents in g, a removal of the link $\overline{ij} \in g$ further splits g into two disconnected subnetworks - one containing i and the other one containing j. Let $D^i_{\overline{ij}}(g)$ and $D^j_{\overline{ij}}(g)$ denote these two subnetworks respectively. Furthermore, let $N^i_{\overline{ij}}(g)$ and $N^j_{\overline{ij}}(g)$ be the sets of agents in these two networks respectively. These notations will later be used to establish the concept of efficient link receiver.

2.6 Modified Networks

$g - ij$ is defined as $g - ij = g \setminus \{ij\}$. That is, $g - ij$ is modified from g by simply removing the link $ij \in g$. Similarly, $g + ij = g \cup \{ij\}$ is the network g modified by adding the link ij. Of course, $g - ij + kl = \left(g \setminus \{ij\}\right) \cup \{kl\}$ is the network g modified by removing the link ij and adding the link kl.

Next, consider two disconnected networks g' and g'' and assume that agents i and j are in g' and g'' respectively, then we define $g' \oplus_{ij} g''$ as the network that results from joining the two networks g' and g'' through the addition of the link ij. That is, $g' \oplus_{ij} g'' = g' \cup g'' \cup \{ij\}$. Note that if $ij \in g$ then $D_{ij}^{i}(g) \oplus_{ij} D_{ij}^{j}(g) = g$.

2.7 Quantity Of Information

Let the value of information that each agent possesses be 1. Let $I_{ij}(g) = \sigma^{d_{ij}(g)}$. That is, I_{ij} is the quantity of information that i receives from j in g. We then define the total information that i receives from every agent in the network as $I_i(g) = \sum_{j \in N} I_{ij}(g) = \sum_{j \in N} \sigma^{d_{ij}(g)}$. Observe that $I_{ii}(g) = 1$.

2.8 Cost Function And The Payoffs

Let $N_i^S(g) \equiv \{j \in N; ij \in g\}$ denote the set of all agents with whom i establishes a link. In most of the literature in the strand of two-way flow model including the pioneering work of Bala and Goyal (2000b) (Proposition 5.5) and recent works of Olaizola and Valenciano (2021) and Hoyer and Jaegher (2023), cost function is assumed to be linear, agent homogeneity in link formation cost is also assumed and the benefit that each agent receives is precisely the total information that he receives. To allow my results to be compared with those of Bala and Goyal (2000b), this note will also adopt these assumptions. This leads to the following payoff;

$$U_i(g) = I_i(g) - n_i^S(g)c = \sum_{j \in N} \sigma^{d_{ij}(g)} - n_i^S(g)c$$

where $n_i^S(g) = \left|N_i^S(g)\right|$.

2.9 Nash Networks

Consider a network g^* such that a strategy of i is $g_i^* \subset g^*$ >. Let $g_{-i}^* = g^* \setminus g_i^*$ so that $g^* = g_i^* \cup g_{-i}^*$. g_i^* is said to be a *best response* of i if $U_i(g^*) \geq U_i(g_i \cup g_{-i}^*)$ for every g_i which is a strategy of i. g^* is said to be a *Nash network* if every agent chooses his best response.

2.10 Efficiency Of A Network

Let $W(g) = \sum_{i=1}^{n} U_i(g)$. A network g *dominates* another network g' if $W(g) \geq W(g')$. A network g is efficient if it dominates every other network. Consider the payoff as in Eq. 1, we can express $W(g)$ as

$W(g)=\sum_{i\in N} I_i(g)-\sum_{i\in N} n_i^S(g)c$. We denote the first term on the right as $\bar{I}(g)=\sum_{i\in N} I_i(g)$ and call it total informational quantity of the network g. Similarly, We denote the second term on the right as $\bar{C}(g)=\sum_{i\in N} n_i^S(g)c$ and call it total cost of the network g.

2.11 Efficient Link Receiver

In a minimally connected network g, consider a link $\overline{xy}\in g$. j' *is superior to* j'' *as a transmitter* with respect to the link $\overline{xy}$ if $j', j''\in N^y_{\overline{xy}}(g)$ and $\bar{I}\left(g-\overline{xy}+\overline{xj'}\right)\geq \bar{I}\left(g-\overline{xy}+\overline{xj''}\right)$. Moreover, j' is said to be an *efficient transmitter* with respect to the link $\overline{xy}$ if j' is superior to every agent in $N^y_{\overline{xy}}(g)$ as a transmitter. A link receiver j is said to be an *efficient link receiver* if j is an efficient transmitter with respect to every link $xy\in g$ such that $y=j$. In a network g, a link sender i is said to be an *efficient link sender* if i is an efficient transmitter with respect to every link $xy\in g$ such that $x=i$.

Remark 1. *Assuming the payoff as in Eq. 1, the following can be said about a minimal efficient network;*

1. *A minimal efficient network is such that every link receiver and every link sender is an efficient transmitter. That is, every agent in the network is an efficient transmitter.*
2. *Since a star is the unique efficient network within the class of nonempty minimal network*[4]*, a star is a unique network such that every link sender and every link receiver is efficient.*

2.12 Best Informed Agent

In a minimally connected network g, Let $M\subset N$ be a minimally connected subset of agents and $i,j\in M$. i is *better informed* than j in the set M if $\sum_{k\in M} I_{ik}(g)\geq \sum_{k\in M} I_{jk}(g)$. If i is better informed than every other agent in the set M, then i is said to be a best-informed agent in the set M. Alternatively, if $M=N^x_{\overline{xy}}(g)$ for a link $\overline{xy}\in g$, we then say that i is better informed than j with respect to the link $\overline{xy}$ and i is *best-informed* with respect to the link $\overline{xy}$.

Observe the following differences between the definitions of best-informed agent and efficient transmitter. The definition of best-informed agent revolves around the informational quantity *received by each individual*, while the definition of efficient transmitter revolves around the total informational quantity *received by all agents* in the network. Despite these differences, it turns out

4 *The result that a star is the unique architecture of nonempty minimal efficient network is stated in Proposition 5.5 of Bala and Goyal (2000b).*

that the identity of an efficient transmitter and best-informed agent is identical. This is proven in the next section.

3. Main Analysis: Proposition 1

I first state a lemma that establishes the threshold level of 'small decay' that guarantees that both efficient network and Nash networks are minimally connected. I then relate the concept of efficient transmitter and the concept of best informed agent in Proposition 1. All proofs are relegated to the Appendix.

Lemma 1. *Let the payoff be as in Equation 1. For any* $c>0$ *and* $n\geq 4$ *there exists a threshold level of decay* $\sigma_K<1$ *such that for all* $\sigma>\sigma_K$ *every nonempty efficient network and every nonempty Nash network is minimally connected.*

Importantly, this Lemma 1 implies that the small decay assumption as used in this note is slightly different from that of De Jaegher and Kamphorst (2015). σ_K above guarantees that all nonempty efficient networks and Nash networks are minimally connected, while the threshold level of decay σ_M as in Lemma 4 of De Jaegher and Kamphorst (2015) only guarantees that all nonempty Nash networks are minimally connected. Such a threshold is necessary since Remark 2 in this note intends to establish a connection between nonempty efficient network and nonempty Nash network, while De Jaegher and Kamphorst (2015) intend to provide the characterization of nonempty Nash networks that is minimally connected[5].

Next, a pivotal lemma that is the basis for Proposition 1 is established below.

Lemma 2. *Let* g' *and* g'' *be two minimally connected networks such that* $g'\cap g''=\varnothing$ *and* N' *and* N'' *be the set of agents in* g' *and* g''

5 Note that the result of Proposition 1 below would still hold even if σ_K is replaced by σ_M as in Lemma 4 of De Jaegher and Kamphorst (2015) since Proposition 1 only points out that every link receiver in a minimally connected Nash network is efficient. Note also that σ_K above is never less than σ_M for any cost c. Intuitively, the existence of σ_M as in De Jaegher and Kamphorst (2015) rests upon the fact that the benefit for a link sender from establish a nonminimal link becomes marginal once the decay is sufficiently small and hence cannot outweigh the additional link formation cost. σ_K above also rests upon the same analogy except that the decay also needs to be sufficiently small to guarantee that the benefit to *all agents* cannot outweigh the additional link formation cost. Finally, note also that because $\sigma_K\geq\sigma_M$ Proposition 1 below is established having in mind that all properties of nonempty minimal Nash networks are as characterized in Proposition 1 of De Jaegher and Kamphorst (2015).

respectively. Let $x \in N'$ *and* $y \in N''$. *Define* $g = g' \oplus_{\overline{xy}} g''$. *Then,*

1. $\sum_{i \in N'} I_i(g) = \overline{I}(g') + \sigma I_x(g') I_y(g'')$
2. $\sum_{i \in N''} I_i(g) = \overline{I}(g'') + \sigma I_x(g') I_y(g'')$
3. *and hence, as a corollary,* $\overline{I}(g) = \overline{I}(g') + \overline{I}(g'') + 2\sigma I_x(g') I_y(g'')$

This Lemma states that the total information in a network - $\overline{I}(g)$ - can be calculated by assuming as if the the network is split via a link $xy \in g$. $\overline{I}(g)$ then depends $\overline{I}(g')$, $\overline{I}(g'')$, $I_x(g')$ and $I_y(g'')$. Thus, if a link $xy \in g$ is replaced by another link xz where z also belongs to the same component as y , then whether total information in the network would improve depends solely on whether $I_z(g'') > I_y(g'')$. This leads to the main result of this note below:

Proposition 1. *Let the payoff be as in Equation 1,* $n \geq 4$ *and let the decay be* $\sigma \in (\sigma_K, 1)$. *In a minimally connected network consider a link* $\overline{xy} \in g$ *,* j' *is superior to* j'' *as a transmitter with respect to the link* $\overline{xy}$ *if and only if* j' *is better informed than* j''. *Consequently, (i)* j' *is an efficient transmitter if and only if* j' *is best informed with respect to the link* $\overline{xy}$, *(ii) in a Nash network every link receiver is an efficient link receiver and (iii) A minimally connected Nash network and a minimally connected efficient network share a similarity: every link receiver is an efficient link receiver*[6].

The proof is given in the Appendix and the intuition is given as follows. Lemma 1 in De Jaegher and Kamphorst (2015) shows that an agent whose network position is 'in the middle' of the other agents tends to be a best informed agent because his position implies that each chain through which information *arrives to him* is relatively short, resulting in him suffering less decay[7]. By the same analogy, being in the middle means that each path through which information *arrives to other agents* from him is also relatively short, resulting in the fact that information that reaches to other agents suffers relatively less decay. In other words, once an agent's position is optimal for receiving information

---

[6] *A straightforward corollary of this Proposition 1 is that in a minimally connected efficient network every link sender and link sender are efficient and, equivalently, best-informed.*

[7] Definition 2 in De Jaegher and Kamphorst (2015) defines a middle agent as follows. "Consider a minimal connected subset of players M, $M \subset N$. We say that player j is in the middle of set M in network g if for each neighbor k of j in network g the following holds: in g more than half of the players in M (including k and j) are closer to j than to k." Lemma 1 in De Jaegher and Kamphorst (2015) then guarantees that a middle agent is always a best-informed agent in a network. More generally, even if there is no agent whose position is in the middle, Lemma 1 in Charoensook (2020) shows that a "positionally optimal" agent always exists.

for his own benefit, it also becomes optimal for transmitting information for the benefits of others. Example 1 below illustrate this intuition.

Example 1. *Let us reconsider our motivating illustration mentioned in the Introduction as in Figure 1. In Figure 1(a), we can see that Susan's position is in the middle of John and Pete. This makes her a unique best-informed agent among the three. Indeed, if we set* $\sigma = 0.99$ *then* $\overline{I}(Susan)1 + 2x0.99 = 2.98 > \overline{I}(John) = \overline{I}(Pete) = 2.97 = 1 + 0.99 + 0.99^2$ *. Moreover, according to our Proposition 1 Susan's position also makes her a superior transmitter of information relative to John and Pete. To understand why, observe that if John establishes a link with Susan as in Figure 1(b), this link yields him the benefits of* 2.98σ *due to the fact that he becomes one link away from Susan. By the same analogy, Matt and Michael receive the benefits of* $2.98\sigma^2$ *due to the fact that they are two links away from Susan.*

On the other hand, if Frank establishes a link with John or Pete, who is less informed than Susan, as in Figure 1(c)? Then from this link establishment Frank receives 2.97σ*, Mike and Matt receive* $2.97\sigma^2$*. Because* $2.97\sigma < 2.98\sigma$ *and* $2.97\sigma^2 < 2.98\sigma^2$*, we conclude that Frank, John and Pete always receive less informational quantity if Frank establishes a link with John or Pete rather than with Susan. Thus, Susan is an efficient transmitter and, according to the previous paragraph, a better informed agent relative to John and Pete. This echoes what our Proposition 1 states "* j' *is superior to* j'' *as a transmitter with respect to the link* $\overline{xy}$ *if and only if* j' *is better informed than* j''*."*

In terms of strategic action, observe also that Frank would also prefer to form a link with Susan rather than John or Pete because he always receives less if the link with Susan is replaced by a link with John or Pete ($2.97\sigma < 2.98\sigma$ *). Thus, Susan is an optimal choice as a link receiver from the strategic point of view of Frank and also, according to the previous paragraph, an efficient transmitter. This echoes what our Proposition 1 states, "...in a Nash network every link receiver is an efficient link receiver."*

Remark 2. *Let the payoff be as in Equation 1,* $n \geq 4$ *and let the decay be* $\sigma \in (\sigma_K, 1)$*, a minimally connected efficient network, which is a star, and a Nash network have the following difference*[8]*. In a minimally connected Nash network a link sender is not necessarily efficient, while it is so in a minimally connected efficient network.*

8 *The result that a star is a unique architecture of nonempty minimally connected efficient network is stated in Proposition 5.5 of Bala and Goyal (2000b).*

4. Applications To The Literature

Finally, I turn to discuss some potential applications of Proposition 1 to the literature. Specifically, a major technical difficulty in the literature is that without decay the set of efficient networks can be very large. This section develops two algorithms that can, in part, solve this problem. Example 2below provides a motivational case in point.

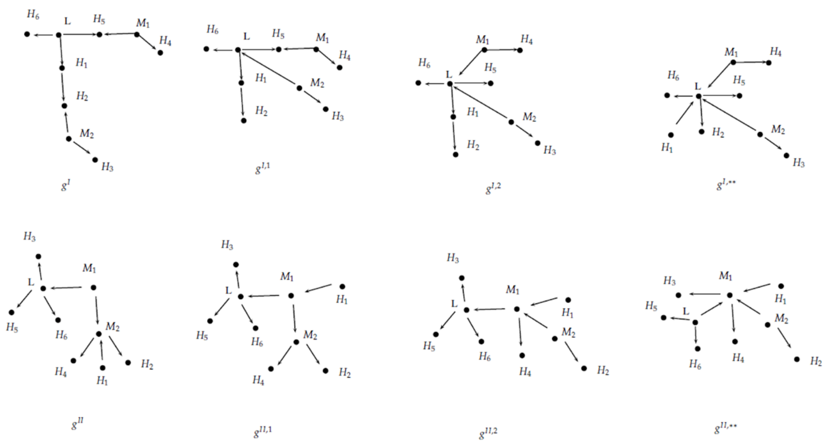
Example 2. *Consider the following example, which is modified from Example 1 in Unlu (2018). Let* $n = 9$ *. Let* c_{ij} *be the link formation cost that* i *pays to establish a link with* j *. Let us assume player cost heterogeneity. That is,* $c_{ij} = c_i$ *for every* $i \neq j$ *. Let* $\sigma = 1$ *(i.e., no decay),* $V = 100$ *where* V *is the value of information possessed by each agent and let the payoff be* $\pi_i(g) = 100\sum_{j\in N}\sigma^{d_{ij}(g)} - \left(n_i^s c_i\right)^2$ [9]*. Let agents consist of 3 groups:* $\{L\}$ *,* $\{M_1, M_2\}$ *,* $\{H_1, H_2, H_3, H_4, H_5, H_6\}$ *and set* $c_L = 1.2$ *,* $c_{M_1} = c_{M_2} = 1.5$ *,* $c_{H_1} = c_{H_2} = c_{H_3} = c_{H_4} = c_{H_5} = 3$ *. As in Unlu (2018) without decay, all networks in Figure 2 are efficient. To see why, observe that all these networks have the same total informational quantity (due to no decay assumption) and total link formation cost since agent* L *sponsors one link, intermediate cost agents* M_1, M_2 *sponsor two links and high cost agents sponsor no links except* H_1 *who sponsors one link*[10]*. Any deviation from such link sponsorship profile will incur a higher total link formation cost due to the convexity of the cost function. For example, if the lowest cost agent* L *sponsors four links instead of three links and the highest cost agent* H_1 *sponsors no link instead of one link, then the total cost of link formation increases from 930.96 to 932.04.*

---

9 *Note that the payoff here assumes a cost function that is more general than that of Equation 1.*

10 *Note that, of course, there are also other networks beside these eight networks that are efficient for the case of no decay. Specifically any network such that* L *sponsors one link, intermediate cost agents* M_1, M_2 *sponsor two links and high cost agents sponsor no links except* H_1 *who sponsors one link is efficient.*

FIGURE 2
EXAMPLE 2



A natural question that arises is how to narrow down such a large set of efficient networks. In what follows, by introducing the small decay assumption, I establish two algorithms that effectively serve this purpose. These two algorithms are simply a straightforward application of our main results as in Proposition 1[11]. while the formal statements of these two algorithms are included in the Appendix, their results are stated below. We first begin with some preliminary notations.

5. Preliminary Notations

Let $n_i^S(g) = |g_i|$. That is, $n_i^S(g)$ is the number of links that agent i sponsors in g. We say that two networks g and g' are $LS-$*equivalent* if $n_i^S(g) = n_i^S(g')$ for every agent $i \in N$[12]. For example, all networks in Figure 2 are $LS-$*equivalent* because, as mentioned in the above Example 2, "agent L sponsors one link, intermediate cost agents M_1, M_2 sponsor two links and high cost agents sponsor no links except H_1 who sponsors one link." Next, we say that a network g' is *improved* from the network g if $I(g') > I(g)$ and g' and g are LS-equivalent. If $I(g') > I(g)$ we say that g' is *superior* to g. If $I(g') \geq I(g)$ for every $g \neq g'$ then we say that g' is *optimal* network.

11 These two algorithms are developed upon the suggestions of an anonymous reviewer, whom I would like to thank.

12 LS stands for link sponsorship.

Next, recall that a chain between i and j in a network g, denoted by $\bar{P}_{ij}(g)$, is defined as a sequence of agents $\{\overline{i_0 i_1},...,\overline{i_{k-1} i_k}\}$ such that $i_0 = i, i_k = j$. A path is defined similarly, except that link sponsorship is one-directional. That is, a path from i to j is a sequence of agents $\{i_0 i_1,...,i_{k-1} i_k\}$ such that $i_0 = i, i_k = j$. A path from i to j in a network g is denoted by $P_{ij}(g)$. In a minimal network, a link ij is said to *point away* from an agent i' if ij is the last link in a chain between i' and j. A path is said to point away from an agent i' if every link in this path points away from i'. By the same analogy, a link ij is said to *point towards* an agent i' if ij is the last link in a chain between i' and i. A path is said to point towards i' if every link in this path points towards i'. Finally, in a minimal network a terminal agent is an agent that has precisely one link.

Next, we define the following sets of links. Let i^* be a best-informed agent in the network g. $L^{S,1}(g) = \{ij \in g \mid ij \text{ points towards } i^* \text{and} i, j \neq i^*\}$ and $L^{S,1}_{\text{MOD}}(g) = \{ii^* \notin g \mid \text{every link } ij \in L^{S,1}(g) \text{ become s} ii^*\}$. For example, in Figure 2 $L^{S,1}(g^I) = \{M_2H_2, M_1H_5\}$ and $L^{S,1}_{MOD}(g) = \{M_2L, M_1L\}$. Next, let $\hat{P}_{it}$ be a maximal multi-link path pointing away from a best -informed agent i^* to a terminal agent t. Let $L^{S,2}(g)$ be the set of all links that belong to such path(s) in g. We partition $L^{S,2}(g)$ into two sets: $L^{S,2^A}(g)$ is the set of all links ij such that ij is the first link on each path $\hat{P}_{it} \in L^{S,2}(g)$ and $L^{S,2^B}(g)$ is the rest of the links in $L^{S,2}(g)$. Two more sets are introduced as modifications of $L^{S,2^A}(g)$ and $L^{S,2^B}(g)$ as follows. $L^{S,2^A}_{MOD}(g) = \{i't \notin g \mid i'j' \in L^{S,2^A}(g) \, and \, i'j' \in \hat{P}_{i't}\}$. That is, for each link in $L^{S,2^A}(g)$ we fix the link sender but changes the link receiver to be a terminal agent. $L^{S,2^B}_{MOD}(g) = \{i'i^* \notin g \mid \text{there is an agent } j' \text{ such that } i'j' \in L^{S,2^B}(g)\}$. That is, for each link in $L^{S,2^B}(g)$ we fix the link sender but changes the link receiver to be the best-informed agent i^*. For example, consider the network $g^{I,2}$ in Figure 3. $L^{S,2}(g^{I,2})$ consists of two links - LH_1, H_1H_2. Hence, $L^{S,2^A}(g^{I,2}) = \{LH_1\}$, $L^{S,2^B}(g^{I,2}) = \{H_1H_2\}$, $L^{S,2^A}_{MOD}(g) = \{LH_2\}$, $L^{S,2^B}_{MOD}(g^{I,2}) = \{H_1L\}$.

We are now ready to state the first result of this section, which shows that Algorithm 1 always leads to an improved network.

Corollary 1. *Let the decay be small such that* $\sigma \in (\sigma_K, 1)$. *Consider a minimally connected network* g *such that* $L^{S,1}(g) \neq \emptyset$. *Let* $ij \in L^{S,1}(g)$. *Then the network* $g' = g - ij + ii^*$ *is improved from* g. *Thus, as in Algorithm 1 the final network* $g^* = (g \setminus L^{S,1}(g)) \cup L^{S,1}_{MOD}(g)$ *where* $L^{S,1}(g^*) = \emptyset$ *is improved from* g.

The proof of this corollary is straightforward and hence is omitted. Intuitively, in a network g Algorithm 1 replaces a link ij with a link ii^* where i^* is a best-informed agent. To understand why this leads to an improved network, observe that because we assume that in g i^* is a best-informed agent we know that in $g-ij$ i^* remains a best-informed agent. Thus, by Proposition 1 i^* is superior as a transmitter compared to i. It follows that $g-ij+ii^*$ is superior to $g-ij+ij=g$. The example below illustrate this intuition.

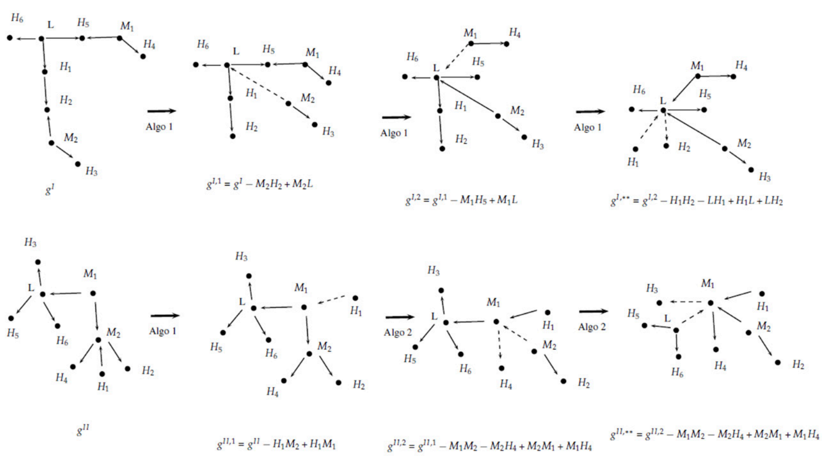
Example 3 *(Example of Algorithm 1)***.** *Figure 3 shows how the network g^{I} is improved into $g^{I,1}$, which is further improved into the final network $g^{I,2}$. It also shows how g^{II} is improved into the final network $g^{II,1}$.*

Next, we establish a result that guarantees that our Algorithm 2 always leads to an improved network. This result is a (non)straightforward corollary of our Proposition 1. The proof is relegated to the Appendix.

Corollary 2. *Let the decay be small such that $\sigma \in (\sigma_K, 1)$). In a minimally connected network g', let there be a two-link path consisting of three agents i', i and t where $it, i'i \in g$ and t is a terminal agent. Then, a modified network $g'' = g' - i'i - it + i't + ii^*$ is improved from g'. Thus, as in Algorithm 2 the final network $g^{**} = \left(g' \setminus L^{S,2}(g')\right) \cup L^{S,2^A}_{MOD}(g') \cup L^{S,2^B}_{MOD}(g')$ is improved from g'. Moreover, $L^{S,2}(g^{**}) = \varnothing$.*

FIGURE 3

AN EXAMPLE OF ALGORITHM 1 AND 2



Example 4. *Consider the network* $g^{I,2}$ *and Figure 3. Through one iteration of Algorithm 2 we achieved the network* $g^{I,**} = g^{I,2} - H_1H_2 - LH_1 + H_1L + LH_2$ *(* $g^{II,**} = g^{II,2} - M_1M_2 - M_2H_4 + M_2M_1 + M_1H_4$ *), and through two iterations of Algorithm 2 the network* $g^{II,1}$ *is transformed into* $g^{II,2}$ *and then* $g^{II,**}$*. Making use of Corollary 2we have* $\bar{I}(g^{I,**}) > \bar{I}(g^{I,2})$ *and* $\bar{I}(g^{II,**}) > \bar{I}(g^{II,2}) > \bar{I}(g^{II,1})$ *.*

Finally, Algorithms 1 and 2 and the resultant Corollary 1 and 2 to establish the main result of this subsection below:

Corollary 3. *Let the decay be small such that* $\sigma \in (\sigma_K, 1)$ *). Any minimal network* g *such that* $L^{S,1}(g) \cup L^{S,2}(g) \neq \varnothing$ *can be improved through Algorithm 1 and 2 above. It can be improved into network* g^{**} *such that* $L^{S,1}(g^{**}) \cup L^{S,2}(g^{**}) = \varnothing$ *through the following procedure. First, use Algorithm 1 to obtain* $g^* = (g \setminus L^{S,1}(g)) \cup L^{S,1}_{MOD}(g)$*. Then use Algorithm 2 to obtain the final network* $g^{**} = (g^* \setminus L^{S,2}(g^*)) \cup L^{S,2^A}_{MOD}(g^*) \cup L^{S,2^B}_{MOD}(g^*)$*. This* g^{**} *has the following properties:*

4. *if an agent* i^* *is a best-informed agent in the initial network* g *then* i^* *is also a best-informed agent in* g^{**}*.*
5. *every chain from* i^* *to a terminal agent has at most two links. If a chain between* i^* *and a terminal agent is precisely one link, then either* i^* *or the terminal agent sponsor the link. If the chain between* i^* *and a terminal agent has precisely two links, then neither* i^* *nor the terminal agent act as a link sender.*
6. *since every chain from* i^* *to a terminal agent has at most two links,* g^{**} *has a diameter of at most* 4*.*
7. *in* g^{**} *there are altogether* $K + n^S_{i^*}(g)$ *agents who are one-link away from* i^**, where* K *is the number of all agents who are link senders except* i^* *in* g*, and the rest of the agents are two-link away from* i^**.*

The proof of this corollary is straightforward and hence is omitted. Example 5 below provide an intuitive illustration for each of these properties.

Example 5. *Consider the network* g^I *(* g^{II} *) in Figure 3. Through two iterations of Algorithm 1 and then one iteration of Algorithm 2 we achieve the network* $g^{I,**}$*. Observe that in* g^I *agent* L *is the middle agent and hence the best-informed agent (See Footnote 7). For each of the iterations, observe that there is one more agent that becomes one link away from* L*. Thus,* L *remains a best-informed agent in the network throughout all iterations. This illustrates property (a) in the above Corollary 3. Next, observe that in* $g^{I,**}$ *each path from* L *to a terminal agent has at most two links. In particular, observe that* *the path from L to H_3 has precisely two links, both of which are sponsored by the agent M_2 who lies between the two agents. This illustrates properties (b) and (c) in the corollary above. Finally, observe that in $g^{I,**}$ there are six agents that are one link away from L . Three of these six (five) agents are accessed by i^* , while the other three agents sponsor themselves. This illustrates properties (d) in the corollary above.*

Importantly, the last property in the above Corollary 3 also necessitates that the final network g^{**} achieved through Algorithm 1 and 2 is not necessarily optimal within the class of all $LS-$ equivalent networks[13]. We state this observation as a remark and then provide an example below.

Remark 3. *Let the decay be small such that $\sigma \in (\sigma_K, 1)$). As a corollary of the property (d) in 3, consider two minimally connected $LS-$ equivalent networks g^1 and g^2 such that $L^{S,1}(g) \cup L^{S,2}(g) \neq \varnothing$. Let $i^{*,1}$ and $i^{*,2}$ be the best-informed agent in g^1 and g^2 respectively and let $g^{1,**}$ and $g^{2,**}$ be the final networks achieved through Algorithm 1 and 2. If $n^S_{i^{*,1}}(g^1) > n^S_{i^{*,2}}(g^2)$ then $g^{1,**}$ is superior to $g^{2,**}$. Thus, within the class of LS-equivalent network, only the final network g^{**} such that the best-informed agent is a largest sponsor is a unique optimal network.*

Example 6 below provides an intuitive illustration of Remark 3above.

Example 6. *Consider networks g^I and g^{II} in Figure 3. Observe that g^I and g^{II} are $LS-$equivalent. Next, observe that L and M_1 are the best-informed agents in g^I and g^{II} respectively but $n^S_L(g^I) = 3 > n^S_{M_1}(g^{II}) = 2$* [14]. *Next, consider the two final networks $g^{I,**}$ and $g^{II,**}$ improved from g^I and g^{II} respectively. Observe that $n^1_L(g^{I,**}) = 6 > n^1_{M_1}(g^{II,**}) = 5$ while $n^2_L(g^{I,**}) = 2 > n^2_{M_1}(g^{II,**}) = 3$. Thus, in $g^{I,**}$ there are more agents who are of distant 2 from each other than in $g^{II,**}$, yet there are less agents who are of distant 3 and 4 from each other than in $g^{II,**}$. We conclude that $g^{I,**}$ is superior to $g^{II,**}$.*

Finally, we apply our Algorithms 1 and 2 to narrow down and precisely identify the set of efficient networks in Example 1 as follows:

---

13 In the sense that there exists another network g^{***} that is $LS-$ equivalent to g^{**} such that $\bar{I}(g^{***}) > \bar{I}(g^{**})$.

14 *For each agent i in a network g, we denote the set of agents that are of k-links away from i as $N^k_i(g)$. Naturally, $n^k_i(g) = |N^k_i(g)|$ is number of agents that are of k-links away from i.*

Example 7. *As a continuation of Example 2, first recall that all networks in Figure 2 are efficient networks if there is no decay. Now, if we assume small decay by setting* $\sigma = 0.999$*, we can partially rank these networks and hence narrow down the set of efficient networks as follows. First, making use of Algorithm 1 and Algorithm 2 as shown in Example 3 and 4 we conclude that* $\bar{I}(g^{I}) < \bar{I}(g^{I,1}) < \bar{I}(g^{I,2}) < \bar{I}(g^{I,**})$ *and* $\bar{I}(g^{II}) < \bar{I}(g^{II,1}) < \bar{I}(g^{II,2}) < \bar{I}(g^{II,**})$*. Second, because we assume player heterogeneity in link formation cost and that all these networks are LS-equivalent, we know that all these networks have the same total link formation cost. Combining these two reasons, we conclude that* g^{II} *is dominated by* $g^{II,1}$*, which is dominated by* $g^{II,2}$*, which is dominated by* $g^{II,**}$*. Thus, among eight networks that are efficient when there is no decay, six are eliminated as candidates for efficient networks if small decay is assumed. Only* $g^{I,**}$ *and* $g^{II,**}$ *remain candidates for an efficient network.*

Next, we answer the natural question that arises: which network $g^{I,**}$ *or* $g^{II,**}$ *- dominates the other? Recall from Example 5that* $\bar{I}(g^{I,**}) > \bar{I}(g^{II,**})$ *. Following the same line of reasoning as in the above paragraph, we conclude that* $g^{I,**}$ *dominates* $g^{II,**}$*. Indeed, recall from Remark 3 that within the class of LS-equivalent network,* $g^{II,**}$ *is a unique optimal network because its best-informed agent* L *is the largest sponsor. Furthermore, recall that we set our decay to be very small, that is,* $\sigma = 0.999$*. Thus, any network that does not dominate* $g^{I,**}$ *in the case of no decay also does not dominate* $g^{I,**}$ *for* $\sigma = 0.999$*. It follows that* $g^{I,**}$ *is a unique efficient network.*

Observe how useful our Algorithms 1 and 2 are. From eight efficient networks in case of no decay as in Example 2, this set is narrowed to just one efficient network by assuming small decay and applying Algorithms 1 and 2.

6. Conclusions

In this note, I show that in a Nash network every link receiver is an efficient transmitter of information in the simple model of two-way flow network with nonrival information pioneered by Bala and Goyal (2000b) if small decay of information is assumed. This results in an insight that a strategic decision of self-interested agents to form links can, in part, leads to a socially desirable outcome. In Section 4, I also show how this result can be applied to refine and resolve the technical difficulties in some existing literature in terms of characterization of efficient networks.

It is important to keep in mind, however, that this note follows the convention in the literature by assuming that the benefit of each agent is precisely the total information he receives in the network. Consequently, how a more general form of benefit function could alter the characteristic of efficient network and the main result of this note becomes a future research question to explore. In addition, another open question is of whether the result of this note can be extended to a more general case such that any level of decay is assumed.

Appendix

Algorithms

Algorithm 1.

1. Initialization:

a. Begin with the network g such that $L^{S,1}(g) \neq \varnothing$.

2. Selection:

a. Choose any link $ij \in L^{S,1}(g)$.

3. Modification:

a. Obtain a modified network g' by performing the following operations on g:

i. Remove the link ij.

ii. Add the link ii^*.

b. Thus, $g' = g - ij + ii^*$. Note that $L^{S,1}(g') = L^{S,1}(g) \setminus \{ij\}$

4. Check and Repeat:

a. If $L^{S,1}(g') \neq \varnothing$:

i. Set $g = g'$.

ii. Repeat steps 2 and 3 with the updated network s.

5. Termination:

a. Continue the process until $L^{S,1}(g') = \varnothing$.

b. Denote the final network as g^*. Note that $g^* = (g \setminus L^{S,1}(g)) \cup L^{S,1}_{MOD}(g)$, where g is the initial network.

Algorithm 2.

1. Initialization:

a. Begin with the network $g^{`}$ such that $L^{S,2}(g^{`}) \neq \varnothing$.

2. Selection:

a. Choose two links $i_0 i_1$ and $i_1 i_2$ in $L^{S,2}(g^{`})$ such that i_2 is a terminal agent.

3. Modification:

a. Obtain a modified network $g^{``}$ by performing the following operations on $g^{`}$:

i. Remove the link $i_0 i_1$ and $i_1 i_2$.

ii. Add the link $i_0 i_2$ and $i_1 i^*$.

b. Thus, $g^{``} = g^{`} - i_0, i_1 - i_1 i_2 + i_0, i_2 + i_1 i^*$.

4. Check and Repeat:

a. If $L^{S,2}(g^{``}) \neq \varnothing$:

i. Set $g^{``} = g^{`}$.

ii. Repeat steps 2 and 3 with the updated network $g^{`}$.

5. Termination:

a. Continue the process until $L^{S,2}\left(g^{`}\right)=\varnothing$.

b. Denote the final network as g^{**}.

USEFUL LEMMATA

Proof 1 (Proof of Lemma 1). *According to Proposition 1 and Lemma 1 in Unlu (2018), in the absence of decay, every nonempty efficient network is minimally connected. Since the aggregate payoff* $W(g)$ *is continuous in* σ *, it follows that for any* $c>0$*, there exists* $\sigma_{M'}<1$ *such that if* $\sigma>\sigma_{M'}$*, then every efficient network is minimally connected.*

Next, recall from Lemma 4 in De Jaegher and Kamphorst (2015) that for any $c>0$ *and* $n\geq 4$*, there exists* $\sigma_M<1$ *such that if* $\sigma>\sigma_M$*, then every Nash network is minimal. Also, recall from Lemmas 5 and 6 in De Jaegher and Kamphorst (2015) that every nonempty minimal Nash network is connected.*

Hence, there exists $\sigma_K<1$*, where* $\sigma_K=\max(\sigma_M,\sigma_{M'})$*, such that if* $\sigma>\sigma_K$*, then every efficient network and every Nash network is minimally connected.*

Proof 2 (Proof of Lemma 2). *First, consider* $i\in N'$*. Observe that in* g' i *receives information of agent in* g'' *via the agent* $x\in N'$ *who, in turn, receives from the* $y\in N''$ *with whom he is one-link away because* $\overline{xy}\in g$*. This fact implies that* $\sum_{l\in N''}I_{xl}(g)=\sigma I_y(g'')$ *and* $\sum_{l\in N''}I_{il}(g)=\sigma^{d_{ix}(g')}\sigma I_y(g'')$*. Next, denote* $\overline{I}_{y\to x}(g)$ *- the total information that all agents in* g' *receives from* g'' *via the link* $\overline{xy}$*. That is,*

$$\overline{I}_{y\to x}(g)=\sum_{i\in N'}\left(\sum_{l\in N''}I_{il}(g)\right)=\sum_{i\in N'}\left(\sigma^{d_{ix}(g')}\sigma I_y(g'')\right)=\sigma I_y(g'')\sum_{i\in N'}\left(\sigma^{d_{ix}(g')}\right)=\sigma I_y(g'')I_x(g').$$

On the other hand, we know that the total information that all agents in g' *exchange with each other is* $\overline{I}(g')$*. This fact and the above expression lead to*

$$\sum_{i\in N'}I_i(g)=\overline{I}(g')+\overline{I}_{y\to x}(g)=\overline{I}(g')+\sigma I_y(g')I_x(g'').$$

This completes the first part of this lemma. The second part of this lemma follows the same analogy so that;

$$\sum_{i\in N''}I_i(g)=\overline{I}(g'')+\sigma I_y(g')I_x(g'').$$

Lastly, observe that part (iii) is simply due to the fact that $\sum_{i\in N}I_i(g)=\sum_{i\in N'}I_i(g)+\sum_{i\in N''}I_i(g)$*. This completes our proof.*

Proof Of Proposition 1

Proof 3. First, by Lemma 1 we know that if decay is sufficiently small then all Nash networks and efficient networks are minimally connected. Let $g' = D^x_{\overline{xy}}, g'' = D^y_{\overline{xy}}, N' = N^x_{\overline{xy}}$ and $N'' = N^y_{\overline{xy}}$ so that $g = g' \oplus_{\overline{xy}} g''$ and $g - \overline{xy} + \overline{xj'} = g' \oplus_{\overline{xj'}} g''$ and $g - \overline{xy} + \overline{xj''} = g' \oplus_{\overline{xj''}} g''$ for any $j', j'' \in N^y_{\overline{xy}}$. Using Equation in (3) of Lemma 2 we know that $\overline{I}\left(g - \overline{xy} + \overline{xj'}\right) \geq \overline{I}\left(g - \overline{xy} + \overline{xj''}\right)$ if and only if $\overline{I}_{j'}\left(g''\right) = \overline{I}_{j'}\left(D^y_{\overline{xy}}\right) \geq \overline{I}_{j''}\left(g''\right) = \overline{I}_{j''}\left(D^y_{\overline{xy}}\right)$. Hence, j' is superior to j'' as a transmitter if and only if j' is better informed than j'' and j' is an efficient transmitter if and only if j' is best informed with respect to the link $\overline{xy}$.

Finally, to prove (ii), recall Remark 1 in De Jaegher and Kamphorst (2015) that if a link $ij \in g$ and g is Nash then j is a best-informed agent with respect to ij.

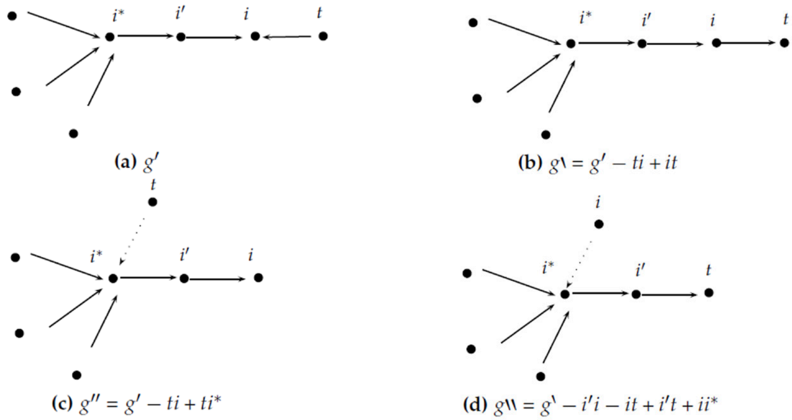
Proof 4 (Proof of Corollary 2). *First, consider the network g' as in Figure 4a, which has i^* as a best-informed agent and a terminal agent t who sponsors himself by accessing another agent i, who receives a link from another agent i'. Observe that if we modify g' into $g^{`} = g' - ti + it$ as in Figure 4b then $g^{`}$ has a two-link path consisting of three agents i', i and t where $it, i'i \in g$ and t is a terminal agent, which is precisely what this lemma assumes. Observe further that by our construction g' and $g^{`}$ share the same architecture. Thus, $\overline{I}(g') = \overline{I}(g^{`})$.*

Next, we modify g' and $g^{`}$ into $g'' = g' - ti + ti^*$ and $g^{``} = g^{`} - i'i - it + i't + ti^*$ as in Figures 4c and 4d. Again, by our construction, $g''$ and $g^{``}$ share the same architecture. Thus, $\overline{I}(g'') = \overline{I}(g^{``})$.

Next, because $g'' = g' - ti + ti^*$ and i^* is a best-informed agent by Proposition 1 we have $\overline{I}(g'') > \overline{I}(g')$. Finally, making use of the last sentences of the above two paragraphs we conclude that $\overline{I}(g'') = \overline{I}(g^{``}) > \overline{I}(g^{`}) = \overline{I}(g')$. This completes our proof.

FIGURE 4

OUR NETWORKS AS IN THE PROOF OF COROLLARY 2



References

Bala, V., and Goyal, S. (2000a). "A Strategic Analysis of Network Reliability," Review of Economic Design, Vol. 5(3); 205–228.

Bala, V., and Goyal, S. (2000b). "A Noncooperative Model of Network Formation," Econometrica, Vol. 68(5); 1181–1229.

Bloch, F., and Dutta, B. (2009). "Communication Networks with Endogenous Link Strength," Games and Economic Behavior, Vol. 66(1); 39–56.

Bloch, F., and Jackson, M.O. (2007). "The Formation of Networks with Transfers Among Players," Journal of Economic Theory, Vol. 133(1); 83–110.

Breitmoser, Y., and Vorjohann, P. (2013). "Efficient Structure of Noisy Communication Networks," Mathematical Social Sciences, Vol. 66(3); 396–409.

Charoensook, B. (2020). "On the Interaction Between Small Decay, Agent Heterogeneity and Diameter of Minimal Strict Nash Networks in Two-Way Flow Model," Annals of Economics and Finance, Vol. 21; 331–361.

De Jaegher, K., and Kamphorst, J. (2015). "Minimal Two-Way Flow Networks with Small Decay," Journal of Economic Behavior & Organization, Vol. 109; 217–239.

Dutta, B., Ghosal, S., and Ray, D. (2005). "Farsighted Network Formation," Journal of Economic Theory, Vol. 122(2); 143–164.

Hojman, D.A., and Szeidl, A. (2008). "Core and Periphery in Networks," Journal of Economic Theory, Vol. 139(1); 295–309.

Hoyer, B., and De Jaegher, K. (2023). "Network Disruption and the Common-Enemy Effect," International Journal of Game Theory, Vol. 52(1); 117–155.

Jackson, M.O. (2008). Social and Economic Networks. Princeton University Press, Princeton, NJ, USA.

Jackson, M.O., and Wolinsky, A. (1996). "A Strategic Model of Social and Economic Networks," Journal of Economic Theory, Vol. 71(1); 44–74.

Olaizola, N., and Valenciano, F. (2021). "Efficiency and Stability in the Connections Model with Heterogeneous Nodes," Journal of Economic Behavior & Organization, Vol. 189; 490–503.

Unlu, E. (2018). "Efficient Networks in Models with Player and Partner Heterogeneity," The Manchester School, Vol. 86(1); 100–118.

Intergenerational Educational Mobility within Chile*

Movilidad Educativa Intergeneracional dentro de Chile

Ercio Muñoz**

Abstract

I provide estimates of intergenerational mobility (IGM) in education at a disaggregated geographic level for Chile, a country with high school-level stratification by socioeconomic status and a decentralized administration of public schools. I document wide variation across municipalities. Relative mobility is correlated to the number of doctors, the number of students per teacher, and earnings inequality. Using a LASSO, I find that the share of students enrolled in public schools, the number of students per teacher, population density, and municipal budget are the strongest predictors of IGM. I also document within-country variability in how parental education is associated with other children's outcomes.

Key words: *Socioeconomic mobility, Geography, Chile, Education.*

JEL Classification: *D63, I24, J62.*

* I am grateful to Jonathan Conning, Christos Giannikos, Roy van der Weide, Núria Rodríguez-Planas, Wim Vijverberg, two anonymous referees, and the participants of the micro brown bag, the dissertation seminar at CUNY and the virtual WEAI 2021 Conference for useful comments. Pablo Vargas and Melanie Saavedra provided outstanding research assistance. Any views expressed in this paper are those of the authors and should not be attributed to the Inter-American Development Bank.

** Inter-American Development Bank. Email: erciom@iadb.org ORCID: https://orcid.org/0000-0003-3327-8431.

DOI: http://dx.doi.org/10.4067/S0718-52862025000200259

Received: January, 2024 Accepted: May, 2025

Resumen

Proporciono estimaciones de movilidad intergeneracional (IGM, por sus siglas en inglés) en educación a un nivel geográfico desagregado para Chile, un país con una fuerte estratificación escolar según el nivel socioeconómico y una administración descentralizada de las escuelas públicas. Documento una amplia variación entre municipios. La movilidad relativa está correlacionada con el número de médicos, la cantidad de estudiantes por docente y la desigualdad de ingresos. Usando un modelo LASSO, encuentro que la proporción de estudiantes matriculados en escuelas públicas, la cantidad de estudiantes por docente, la densidad poblacional y el presupuesto municipal son los predictores más fuertes de la IGM. También documento variaciones dentro del país en la asociación entre la educación de los padres y otros resultados de los hijos.

Palabras clave: *Movilidad socioeconómica, Geografía, Chile, Educación.*

Clasificación JEL: *D63, I24, J62.*

1. Introduction

How much of an individual's educational achievement is due to his or her parents' educational achievements? High persistence in educational outcomes across generations can lead to unrealized human capital potential and inefficient allocation of resources and talents that result in lower economic growth. Moreover, it can be a mechanism by which economic advantage is inherited, as education is linked to the capacity to generate income and wealth. Economists have made important progress in documenting the level of intergenerational mobility (IGM) in education (i.e., the relationship between educational outcomes of parents and children) for many countries (see Van der Weide et al. 2024).[1] However, the evidence at the country level can hide important variation within countries, as it has been shown by a growing literature (for example, Alesina et al. 2021, 2023, Asher et al. 2024, Card et al. 2022, Derenoncourt 2022, Hilger 2016, Feigenbaum 2018, Munoz 2024) that estimates IGM for small geographical units, extending the literature on IGM in income initiated by Chetty et al. (2014).

In this paper, I contribute to this literature in three ways. First, I estimate intergenerational mobility in education in Chile at the country, region, and municipality level using census data for a cohort born in the 1990s. I offer eight indicators that describe the association between children's and parents' years

[1] See Torche (2021) for a survey focused on developing countries.

of schooling,[2] capturing different policy-relevant concerns and aspects of this association (e.g., how much more educated children with parents with an extra year of schooling tend to be, how likely are children from low-educated parents to be low-educated, how likely are children to surpass the education of their parents, etc.), therefore providing a broad view of IGM.[3] I provide these estimates in an online data appendix for future research. Second, I show how other children's outcomes, such as teenage pregnancy and tertiary education attendance, are also associated with parental education at the country level and display wide variation within the country. Finally, I explore how the estimates of educational IGM are correlated with a rich set of variables related to income, geography, education, municipal budget, and other characteristics of the municipalities. Furthermore, I investigate by means of a lasso (least absolute shrinkage and selection operator), what correlates have the most predictive power over IGM at the level of municipality.

IGM literature for Chile. Previous studies have used different household and opinion surveys (see for example, Torche 2005, Hertz et al. 2007, Nunez & Miranda 2010, Narayan et al. 2018, Celhay et al. 2010, Celhay & Gallegos 2015, Sapelli 2016, Neidhöfer et al. 2018, Van der Weide et al. 2024, Celhay & Gallegos 2025) to document IGM in income, education, and other socioeconomic measures. However, they all have in common that the samples are not representative at the municipality level, so they focus on country-level estimates. Two exceptions are Celhay & Gallegos (2015) which also explores mobility at the regional level (the coarser administrative unit in which the country is divided), and Cortés Orihuela et al. (2023), which uses labor earnings in the formal sector from administrative records to estimate income mobility at the regional level and between municipalities (the smallest administrative unit) in the Metropolitan Region. More recently, Celhay & Gallegos (2025) analyze intergenerational mobility in education across three generations in six Latin American countries, including Chile. Their work highlights the relevance of multigenerational persistence and provides new evidence on long-term mobility in the region.

Institutional background. Chile is an interesting case study to analyze IGM at the sub-national level. On the one hand, the country is one of the richest economies in the Latin American region and has shown significant progress in poverty reduction and income per capita growth in the last three decades. On the other hand, income inequality is relatively high for OECD standards, and

---

2 Throughout the paper, I will refer to the cohort of interest as "children," and I will refer to their parents or older relatives living in the same household as "parents". I will precisely define who will be considered as a parent in Section II.

3 Deutscher & Mazumder (2023) provide a framework that highlights the key concepts and properties of different indicators of IGM.

previous research has documented high school-level stratification by socioeconomic status (Mizala & Torche 2012, Gutiérrez & Carrasco 2021), which has fueled some educational reforms in the last decade. In addition, the country is marked by the free-market reforms inherited from the military dictatorship (1973-1990). This includes a universal voucher system and decentralization of the administration of public schools, which are managed by municipalities.[4]

In terms of IGM at the country level, the best evidence available at a global scale (Narayan et al. 2018, Van der Weide et al. 2024) shows some interesting findings for Chile. Among the 148 countries for which there are estimates of educational mobility for the cohort born in the 1980s, the country ranks relatively low when a summary statistic of relative mobility, such as one minus the Pearson correlation coefficient between years of schooling of children and parents is used but somewhat more mobile according to one minus the regression coefficient between these two variables.[5] In contrast, the country seems much more mobile when we look at a measure of absolute mobility like the share of students with higher education than their parents. However, when a measure that aims to capture directional mobility from the bottom to the top is considered (i.e., "rags to riches" or poverty to privilege rate as named in Narayan et al. 2018), then the country appears among the least mobile ones (see Figure A1 in the Appendix).[6]

The evolution across different cohorts for these indicators also shows some interesting patterns when compared to simple averages by region as classified in Narayan et al. (2018).[7] Chile does not show much progress in most of the indicators relative to regional averages, except for absolute mobility (share of students with higher education than their parents) and relative mobility measured as $1-\beta$. In contrast, relative mobility measured as $1-\rho$ (independent of the marginal distributions of education) has remained at lower levels than all the regional averages for all the cohorts in the same way as the poverty to privilege ratio (or rags to riches).

4 A recent reform started a process of centralization in 2018.

5 The correlation coefficient can be transformed into the regression coefficient by multiplying it by the ratio of the standard deviation of child schooling to parent schooling. Therefore, differences between them are explained by changes in inequality across generations.

6 Ranked 138 among the 148 available estimates.

7 Figure A2 in the Appendix plots all these indicators across cohorts.

2. Data And Methods

Data. I use individual-level data from the 2017 census of housing and population obtained from the National Institute of Statistics.[8] This statistical operation, which aimed to capture the total population of Chile, includes demographic details such as age, sex, education, household composition, as well as detailed geographical information.

Sample definition. The full-count census database contains information about 17,574,003 individuals. I keep people born in Chile aged between 21 and 25 years and drop those considered domestic service, living in collective housing, or in transit, which reduces the sample to 1,155,207 individuals, 568,231 men and 586,976 women.[9]

Education. The census data contains a variable reporting schooling, regardless of the track or kind of study. When I study how the educational attainment of children relates to the attainment of parents, I take the highest attainment among the individuals in the older generation.[10] Given the typical educational path in Chile, where students start first grade at the age of 6, the average student would be able to attain at most 15 years of schooling at the age of 21. To accommodate for this, the indicators are computed using years of schooling censored at 15 for both children and parents.[11]

Geography. Chile is divided into 16 regions, 56 provinces, and 346 communes or municipalities.[12] The data set contains information on where the interview was conducted and the place of birth in terms of these three administrative divisions. I use the latter to assign people to places and estimate IGM for the country, by region, and by municipality.

Linking individuals across generations. The data set enumerates individuals into households and contains a variable that describes the relationship of each individual with the head of the household. I use this variable to link individuals with their parents or older relatives according to Table 1. In addi-

---

8 The data can be accessed at https://www.ine.gob.cl/estadisticas/sociales/censos-de-poblacion-y-vivienda/censo-de-poblacion-y-vivienda.

9 Van der Weide et al. (2024) use the same age range to estimate IGM in education using survey data for 39 countries.

10 The results are qualitatively similar if I use the average rounded to the nearest integer instead of the maximum.

11 Similar censoring of years of schooling is used in Neidhöfer et al. (2018) with survey data to compute IGM at the country level for 18 countries in Latin America. Figure A3 displays the distribution of educational attainment of parents and children.

12 Chile does not have a commonly used designation of commuting zones, such as the one used in Chetty et al. (2014) for the United States. Many municipalities are within commuting distance in the country, particularly within the Metropolitan area. However, the estimates at the regional level for this specific case are close to what could be considered a commuting zone.

tion, those living only with individuals not identified in the table are matched with other relatives, provided that these relatives are at least 15 years but less than 40 years older than them. In the end, I am able to match approximately 73% of the target sample using specific relationships to the head and an extra 6% using other relatives, reaching a final sample of 833,107 individuals (i.e., a coresidence rate of 79%).[13]

TABLE 1

RELATIONSHIP TO HOUSEHOLD HEAD AND IDENTIFICATION OF DIFFERENT GENERATIONS

Relationship to the head	**Generation**	**Relationship to the head**	**Generation**
Grandparent	-2	Sibling	0
Parent	-1	Sibling-in-law	0
Parent-in-law	-1	Child	1
Head	0	Child-in-law	1
Spouse	0	Spouse/partner of child	1
Legal live-in partner	0	Grandchild	2
Partner	0	Others	Missing

Notes: Categories left missing are: Other relative, non-relative, domestic employee, person in collective housing, visitor, and homeless person.

Coresident sample and potential biases. The use of a sample that only includes individuals cohabiting with their parents is a relatively standard approach in the literature that uses census data and linked generations (see for example, Alesina et al. 2023, 2021, Card et al. 2022, Derenoncourt 2022, Dodin et al. 2024, Feigenbaum 2018, Abramitzky et al. 2021, Ager & Boustan 2021, Munoz 2024).[14] However, there is a potential concern that it may lead to bias in the estimates of IGM as individuals who reside with their parents may systematically differ from those not residing with them (see Emran et al. 2018, Francesconi & Nicoletti 2006).

[13] This closely follows the approach used in Alesina et al. (2021) to link generations with census data from Africa. In comparison, they link 69% of individuals aged 14-25 using specific relationships and 23.6% based on age.

[14] Recent work with survey data have also relied on coresident samples for a large number of countries (see Van der Weide et al. 2024).

Munoz & Siravegna (2023) show evidence suggesting that the bias from the coresidence restriction is relatively small for estimates of some indicators of mobility that use census data and in cases where the bias is larger, there is a low level of re-ranking when these estimates are used to rank economies across time and space by level of mobility relative to the ranking obtained with estimates that use retrospective information (i.e., surveys that ask all individuals for the level of education of their parents). Additional exercises with survey data are also done in Van der Weide et al. (2024), showing that the bias does not generate meaningful re-rankings.

To explore to what extent the estimates reported in the paper are affected by this issue, I compare my estimates of relative mobility (based on the regression coefficient as well as the one based on the correlation coefficient) at the country level with those obtained from recent literature that uses survey data with retrospective information. The estimates are very close, which suggests that the bias is negligible for this particular sample (Figure A4 in the Appendix). Moreover, I explore whether the average coresidence rate at the municipality level is associated with the level of intergenerational mobility and find null to negligible associations (Figure A5 in the Appendix). Similarly, I do not find evidence suggesting that the level of coresidence varies with the level of schooling reported by the children (Figure A6 in the Appendix).[15]

Measurement. I consider eight different indicators that relate to different aspects of educational IGM and for which the choice among them can be justified by the purpose of the analysis (Mazumder 2016, Corak 2020, Deutscher & Mazumder 2023). The first two are derived from a simple OLS regression that relates the educational attainment of children to the attainment of parents. Hence, these measures come from the following specification by municipality c (or country or region):

(1) $$y_{ic}^{y} = \alpha_c + \beta_c y_{ic}^{o} + \epsilon_{ic}$$

Where y_c^y is the educational attainment of individual i (using a sample of individuals with ages between 21-25), y_{ic}^{o} is the attainment of his/her parents or older relatives cohabiting in the same household, and the parameters of interest α_c and β_c are respectively used to measure absolute and relative mobility $(1-\beta_c)$ for municipality of birth c (see Narayan et al. 2018, Torche 2021, for a discussion about the concepts of absolute and relative mobility in education). Given that the expected years of schooling of an individual according to equation 1 depends on the average years of schooling of parents in his/her municipality (in addition to the parameters α_c and β_c), I also compute average

15 Coresidence rates decrease monotonically with age (see Figure A7 in the Appendix).

years of schooling of parents by municipality as the third indicator. The fourth measure relates to the concept of absolute mobility measured as the share of children attaining more years of schooling than their parents (including ties at 15). The fifth measure corresponds to Pearson's correlation coefficient between years of schooling of children and parents, which, in contrast to the regression coefficient, is not affected by the marginal distributions of educational attainment of parents and children. The last three indicators address directional mobility. First, upward IGM (or "rags to riches") is measured as the probability of children reaching the top quintile in the distribution of educational attainment of children in the country (approximately 15 years of schooling) if their parents were in the bottom quintile of educational attainment (approximately less than 10 years of schooling) of parents in the country.[16] Second, intergenerational low education is the probability of attainment in the bottom quintile of the children's distribution (approximately less than 12 years of schooling) when their parent's attainment is also in the bottom quintile of the parent's distribution (approximately less than 10 years of schooling). Finally, intergenerational high education, which is the probability of children's attainment in the top quintile (approximately more than 14 years of schooling) when their parents' attainment is in the top quintile (approximately more than 13 years of schooling).[17] The indicators are summarized in Table 2.

These measures of intergenerational mobility capture different aspects of the association between children's and parents' educational outcomes. The first two indicators (absolute mobility α and relative mobility $1-\beta$) are derived from regression specifications and measure different dimensions of mobility: α represents the expected years of schooling for children whose parents have zero education, providing an anchor for absolute mobility, while $1-\beta$ reflects the degree to which parental education does not determine children's outcomes, with higher values indicating greater mobility. The average parental education ($\bar{Y}$) offers contextual information about the general educational level within each geographic unit. The proportion of children with more schooling than their parents ($\overline{y^{\geq}}$) is another absolute measure, reflecting upward movement regardless of position in the distribution. In contrast, $1-\rho$ is a pure relative mobility measure based on correlation that is invariant to changes in the marginal distributions of education. The final three indicators capture directional mobility through conditional probabilities: rags to riches $P_{1,5}$ measures upward mobility from the bottom to the top of the distribution; intergenerational low $P_{1,1}$ captures persistence at the bottom; and intergenerational high $P_{5,5}$ reflects

[16] The quintiles are defined by sorting individuals by attainment and solving ties randomly.

[17] I also compute these three indicators using quintiles of the distribution of educational attainment within the region or municipality instead of the country. I compare both alternatives in the Appendix.

persistence at the top. Together, these measures provide a comprehensive picture of intergenerational mobility, combining absolute benchmarks, relative positioning, and movement across key points in the educational distribution.

TABLE 2
INDICATORS OF EDUCATIONAL INTERGENERATIONAL MOBILITY

Indicator		**Description**
Absolute mobility	α	OLS estimate of intercept in Eq. 1
Relative mobility (regression coefficient)	$1-\beta$	OLS estimate of slope in Eq. 1
Average education	$\bar{Y}$	Average years of schooling of parents
Above parent	$\overline{y^{\geq}}$	Share with higher schooling than parents
Relative mobility (correlation coefficient)	$1-\rho$	Pearson correlation coefficient
Rags to riches	$P_{1,5}$	Probability of top education conditional on parents in the bottom
Intergenerational low	$P_{1,1}$	Probability of bottom education conditional on parents in the bottom
Intergenerational high	$P_{5,5}$	Probability of top education conditional on parents in the top

Notes: Above parent considers ties at the maximum number of years of schooling in the data as children having higher education than parents. The subscripts in the last 3 rows refer to quintiles. Top and bottom refers to top quintile and bottom quintile.

3. Estimates Of Intergenerational Mobility

In this section I document the level of IGM in Chile derived from my estimates. First, I go over country-level estimates for the eight indicators of IGM described in the previous section. I explore whether there is some evidence of heterogeneity by gender, urban status, and Indigenous population in absolute and relative mobility, and then I go over the estimates of mobility using other outcomes. Second, I document within-country mobility at the region level using the same eight indicators, describe and map the estimates at the municipality level, analyze the correlation patterns between these indicators, and finally explore within-country variation in the effect of parental education on alternative outcomes.

3.1 Country-Level Estimates

I first estimate intergenerational mobility in education at the country level[18], and then I explore some potential heterogeneity across sub-populations such as male versus female, urban versus rural, and Indigenous versus non-Indigenous people in some of the indicators. Moreover, I estimate the association between parental educational attainment and other children's outcomes, such as attending tertiary education and having a child while a teenager in the case of women.

IGM in education. Table 3 summarizes the level of educational IGM using the previously described indicators estimated at the country level with a sample that includes only children between the ages of 21 and 25. The most recent estimates of IGM (at least for a few of these indicators) at the country level available in the literature for Chile are for the cohort born in the 1980s and 1992-1995. Compared to the latter, I find slightly lower relative mobility as measured by $1-\beta$[19] but practically the same level when measured with the $1-\rho$ (0.64 vs. 0.62).[20] In addition, although not constructed in the same way, the indicators of directional mobility ($P_{1,5}$, $P_{1,1}$, and $P_{5,5}$) show a consistent picture with respect to Narayan et al. (2018)'s results in terms of high-persistence at the top of the educational distribution and relatively low chances of reaching the top conditional on having parents at the bottom.

18 Figure A3 in the Appendix displays the distribution of years of schooling for children and parents.

19 For several cohorts, this difference is smaller than the discrepancy between mobility estimated using Latinobarometro and CASEN survey, as reported in Neidhöfer et al. (2018).

20 Figure A4 in the Appendix shows the evolution over time of these indicators in the literature versus my estimates.

TABLE 3

IGM AT COUNTRY-LEVEL

Absolute mobility	α	9.576	Relative mobility	$1-\beta$	0.714
Average education	$\bar{Y}$	11.125	Above parents	$\overline{y^{\geq}}$	0.666
Relative mobility	$1-\rho$	0.642	Rags to riches	$P_{1,5}$	0.088
Intergenerational low	$P_{1,1}$	0.366	Intergenerational high	$P_{5,5}$	0.354

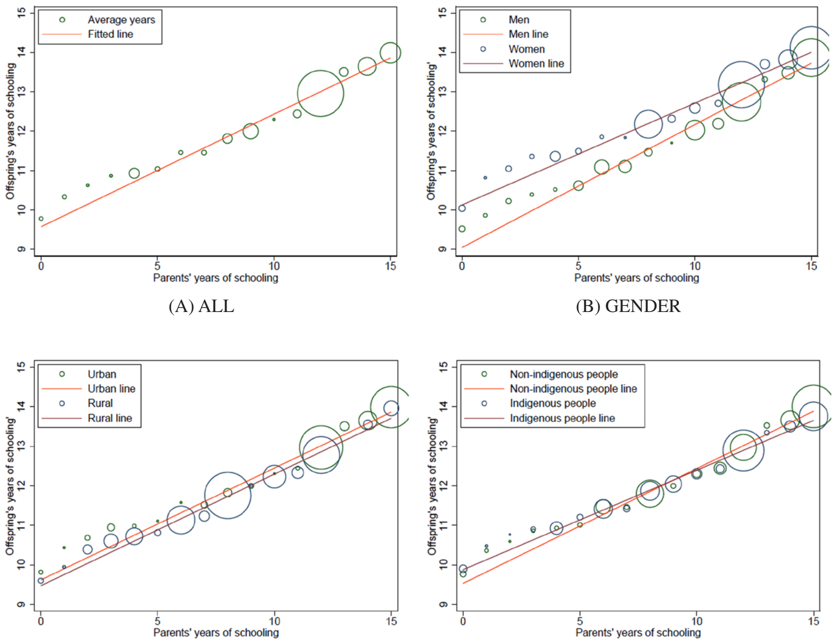
Notes: The table reports estimates of IGM (as described in Table 2) using a sample of individuals aged between 21 and 25 linked to parents or older relatives, as explained in section II.

Figure 1a displays the average attainment conditional on parental education attainment, the relationship appears linear with a deviation only in the lowest level of parental education.[21] When this regression is estimated using sub-populations, I find higher absolute and relative mobility for women compared to men (see Figure 1b). In contrast, I do not find significant differences between rural and urban populations (see Figure 1c), and between Indigenous versus non-Indigenous populations (see Figure 1d). Nonetheless, this does not imply that the expected educational attainment between individuals in urban/rural or Indigenous/non- Indigenous is the same, as can be inferred by the differences in the marginal distributions of parental educational attainment. For example, the number of parents with at least 12 years of education is greater for urban (as well as for non-Indigenous) than rural (and respectively Indigenous population) populations (i.e., the size of the bubbles in Figure 1 is bigger). Table A2 in the Appendix reports the eight indicators computed by subgroup, confirming these findings and highlighting some other differences between groups in other indicators.

[21] Figure A8 in the Appendix displays the transition matrix between children and parental years of schooling, each of them divided into quintiles according to their respective distribution of years of schooling.

FIGURE 1

COUNTRY-LEVEL EDUCATIONAL IGM



(A) ALL (B) GENDER

(C) URBAN VS. RURAL POPULATION (D) INDIGENOUS VS. NON-INDIGENOUS

Notes: The graphs display the average years of schooling of children for each level of schooling of the generation above (highest years of schooling among parents and older relatives living in the same household). The sample includes only individuals between the ages of 21 and 25. The size of the bubble varies according to the number of individuals.

Other children's outcomes. I follow the seminal paper by Chetty et al. (2014) and analyze how family background is associated with two additional children's outcomes: the likelihood of attending tertiary education and the likelihood of having a child while a teenager in the case of women.[22] These outcomes are potentially very consequential in life trajectories, as the first one is positively associated with earnings and other indicators of well-being (Oreopoulos & Petronijevic 2013), while the second is negatively associated with income and some indicators of well-being (Fletcher & Wolfe 2009). Furthermore, these outcomes can be measured at earlier ages than our main outcome (i.e., years of schooling), reducing the magnitude of any potential coresidence bias, as coresidence rates decrease with age.

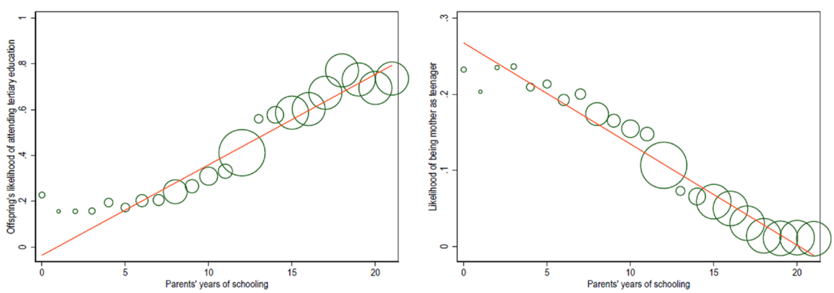
[22] I use the same econometric specification as in Equation 1 with a different dependent variable.

First, I estimate the probability of attending at least one year of tertiary education using a sample of individuals between the ages of 19 and 21. Figure A4c shows this likelihood for each parental educational attainment, finding a positive slope approximately equal to 0.046 with a somewhat prominent discontinuity at 12 years of schooling and a somewhat nonlinear relationship for low values of parents' years of schooling. This contrasts with the virtually linear relationship between parental income rank and college attendance documented for the US in Chetty et al. (2014). Despite these differences and other differences in terms of measurement and concepts, I find similar gaps. The gap in the likelihood of attending tertiary education for individuals with low-educated vs. highly-educated parents is around 60 percentage points while Chetty et al. (2014) documented a gap of 67.5 percentage points in the US for individuals with lowest-income vs. highest-income parents.

Second, I estimate the probability of becoming a mother as a teenager, defined as having a child for females between the ages of 15 and 19. Figure 2b shows this likelihood for each parental educational attainment, finding a negative relationship close to linear with a slope of -0.017. The gap between highly-educated and low-educated parents is around 20-25 percentage points (Chetty et al. 2014, documents a gap of 29.8 percentage points for highest-lowest parents' incomes).

FIGURE 2

OTHER CHILD'S OUTCOMES



(A) TERTIARY EDUCATION ATTENDANCE (B) TEENAGE BIRTH (FEMALES ONLY)

Notes: The first plot displays the likelihood of completing at least one year of tertiary education for each level of education of the generation above (highest years of schooling among parents and older relatives living in the same household). The second plot displays the likelihood of having a child as teenager for each level of education of the generation above. The samples include individuals with age between 19 and 21 (left) and 15 and 19 (right). The size of the bubble varies according to the number of individuals.

3.2 Intergenerational Mobility Within Chile

Region-level estimates. Before presenting the most disaggregated estimates, Table 4 summarizes the eight measures of interest estimated for the 16 regions of Chile. Non-negligible differences can be found across regions in most of these dimensions. For example, the chances of reaching the top quintile of the educational distribution for children with parents at the bottom quintile (i.e., $P_{1,5}$) is more than 200% higher in the northern Arica y Parinacota region relative to Aysén region. Similarly, in terms of absolute mobility (i.e., α), there are regions with more than one year of difference, and relative mobility (i.e., $1-\beta$) is 17% higher in Arica y Parinacota than in Metropolitana de Santiago or Los Rios. When I consider relative mobility measured with the correlation coefficient ($1-\rho$), the level in Arica y Parinacota is approximately 30% higher than in the region with the lowest value (Araucanía).[23]

Municipality-level estimates. I document wide within-country variation. Relative mobility measured as $1-\beta$, excluding places with less than 50 individuals[24], ranges between 0.54 in Quemchi, a municipality located in the south of the country, and 0.97 in San Pedro de Atacama, a municipality located in the north. Non-negligible variation is found in all the indicators studied. Figure A10 in the Appendix shows the distributions of the municipality-level estimates for the eight measures and Table A3 of the Appendix similarly reports some descriptive statistics of these estimates. For all the indicators I can find municipalities with levels at least 100% greater than others, in some cases several times greater.

The measures of mobility based on conditional probabilities derived from quintiles of educational attainment are constructed using the distribution of attainment at the country level for children and similarly for parents. Similar measures could be constructed using the distribution of attainment by municipality. In this case, moving from the bottom to the top may require a higher number of years of schooling in some places compared to others and capture a different aspect of mobility. As an additional exercise, I compute those measures and find that $P_{1,5}$ measures constructed in both ways are highly correlated while $P_{1,1}$ is to a lesser degree while in contrast, $P_{5,5}$ is not correlated (see Figure A11 in the Appendix.).

---

23 Table A4 in the Appendix reports the last three indicators of IGM using the distribution of educational attainment at the country level versus at the region level. There is heterogeneity in the direction of change, but for the three indicators, the range of variation decreases using the local distribution.

24 Figure A9 in the Appendix displays the CDF of the sample size by municipality, showing that less than 5% of the municipalities have less than 50 observations.

Figure 3 maps relative mobility $(1-\beta)$ across the country. There are some regions with clusters of municipalities showing relatively similar levels of IGM, such as the northern regions and more heterogeneity in the center of the country. Figure A15 in the Appendix plots relative mobility dividing the map of the country into three parts, a northern region less the metropolitan region, the metropolitan region, and a southern region. These three regions have municipalities with relatively low and high levels of intergenerational educational mobility. However, in this map the variety in IGM levels of the metropolitan region (where the highest share of the population lives) can be appreciated in more detail.

TABLE 4

REGION-LEVEL ESTIMATES OF IGM STATISTICS

Region	α	$1-\beta$	$\bar{Y}$	$\overline{y^{\geq}}$	$1-\rho$	$P_{1,5}$	$P_{1,1}$	$P_{5,5}$
Tarapacá	9.66	0.74	11.53	0.60	0.70	0.10	0.37	0.32
Antofagasta	9.25	0.71	11.61	0.57	0.68	0.08	0.40	0.31
Atacama	9.54	0.73	10.99	0.62	0.68	0.07	0.40	0.29
Coquimbo	9.44	0.72	10.53	0.65	0.66	0.07	0.38	0.32
Valparaíso	9.61	0.72	11.23	0.65	0.68	0.09	0.35	0.34
Libertador General Bernardo O'Higgins	10.06	0.76	9.95	0.71	0.71	0.10	0.34	0.31
Maule	9.75	0.73	9.84	0.73	0.67	0.09	0.36	0.32
Biobío	10.12	0.74	10.65	0.74	0.66	0.11	0.33	0.37
Araucanía	9.58	0.71	9.88	0.75	0.61	0.07	0.38	0.37
Los Lagos	9.35	0.71	9.77	0.71	0.65	0.07	0.41	0.31
Aysén del General Carlos Ibáñez del Campo	9.38	0.76	9.59	0.65	0.73	0.05	0.44	0.23
Magallanes y de la Antártica Chilena	10.36	0.77	11.33	0.66	0.72	0.10	0.30	0.30
Metropolitana de Santiago	9.38	0.70	11.33	0.64	0.65	0.09	0.37	0.36
Los Ríos	9.46	0.70	10.18	0.72	0.62	0.06	0.38	0.34
Arica y Parinacota	10.76	0.82	11.49	0.61	0.79	0.14	0.28	0.31
Ñuble	10.02	0.74	9.86	0.76	0.68	0.11	0.33	0.36

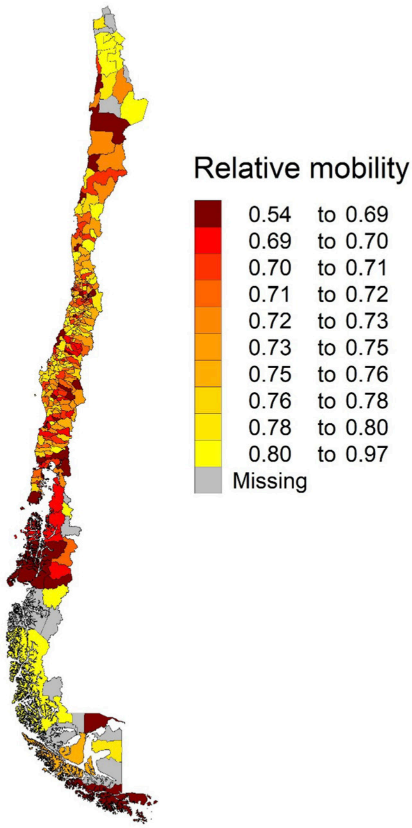
Notes: The table reports region-level estimates of absolute mobility, relative mobility $(1-\beta)$, average parents' education, the share of children with higher education than parents, relative mobility $(1-\rho)$, rags to riches, intergenerational low, and intergenerational high, respectively. A description of the measures can be found in Table 2. Rows are sorted by the official designated number that each region used to have until 2018.

Correlations among different measures of IGM. Table 5 presents the Pearson correlation coefficients between the eight mobility statistics computed at the municipality level. I find the strongest positive correlation to be between absolute and relative mobility, both measured with $1-\beta$ and $1-\rho$. These three measures are at the same time positively correlated to above parents and rags to riches, especially absolute mobility. Intergenerational low is negatively correlated with the other six indicators.

FIGURE 3

INTERGENERATIONAL EDUCATIONAL $(1-\beta)$

MOBILITY WITHIN CHILE



Notes: The map plots relative IGM measured as one minus the regression coefficient (by municipality) between children's years of schooling (using individuals aged between 21 and 25) against parents' years of schooling. Educational attainment is censored at 15. Municipalities with less than 50 individuals are left as missing (Figure A9 in the Appendix displays the CDF of the sample size by municipality). Figure A15 in the Appendix displays a version of the map dividing Chile into three areas.

TABLE 5
CORRELATION AMONG IGM STATISTICS

	α	$1-\beta$	$\bar{Y}$	$\overline{y^{\geq}}$	$1-\rho$	$P_{1,5}$	$P_{1,1}$	$P_{5,5}$
Absolute mobility (α)	1							
Relative mobility ($1-\beta$)	0.912***	1						
Average education ($\bar{Y}$)	-0.0175	-0.146**	1					
Above parents ($\overline{y^{\geq}}$)	0.268***	0.139*	-0.716***	1				
Relative mobility ($1-\rho$)	0.713***	0.874***	-0.0917	-0.0128	1			
Rags to riches ($P_{1,5}$)	0.478***	0.259***	0.296***	0.0604	0.228***	1		
Intergenerational low ($P_{1,1}$)	-0.730***	-0.517***	-0.207***	-0.166**	-0.380***	-0.537***	1	
Intergenerational high ($P_{5,5}$)	-0.00472	-0.141*	0.0369	0.233***	-0.140*	0.236***	-0.0715	1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Other outcomes within Chile. I estimate the relationship between parental education and the two alternative outcomes described in the previous section: the likelihood of attending at least one year of tertiary education, and the likelihood of being a mother as a teenager for females.

Table 6 reports these estimates at the regional-level. There is significant variation across regions in the effect of an additional year of parental schooling on the chances of attending tertiary education. Araucanía shows the strongest effect (0.044), which suggests that the gap between individuals with uneducated parents and those with highly educated ones (21 years) in the chances of attending tertiary education is approximately 92 percentage points (21 × 0.044). A caveat to note is that this calculation may overestimate the effect in light of the non-linearity observed at the national level in Figure A4c for lower levels of parental education. If I assume that the effect is null in the first 5 years of education, then the gap is approximately 70 percentage points. On the other extreme, Aysén region shows the smallest average effect (0.019).

Similarly, the effect of an extra year of parents' schooling on teenage birth rates varies significantly across regions. The effect of one year goes from a fall in the likelihood of a teenage birth equal to 0.8 percentage points in Ñuble to 1.6 percentage in Antofagasta or Coquimbo. This last effect implies a gap between uneducated and highly educated parents of approximately 33.6 percentage points, which again is meaningful but may be an overestimation due to non-linearities.

4. Correlates Of IGM Within Chile

In this section, I examine whether intergenerational educational mobility at the municipality level is correlated with a broad set of variables, including income distribution, educational characteristics, municipal budgets, geographic factors, and other local attributes. Understanding these relationships is crucial to identifying the main factors that contribute to the persistence of educational outcomes across generations. Given the large number of potential predictors, the analysis follows a two-step approach. First, I investigate the correlations between IGM and selected key variables that the literature and theoretical frameworks suggest as particularly relevant. Second, I apply a LASSO (Least Absolute Shrinkage and Selection Operator) regression to identify the strongest predictors of mobility. This method performs a selection of correlates based on their predictive strength.

TABLE 6

PARENTAL EDUCATION EFFECT ON OTHER OUTCOMES

Region	**Tertiary education**	**Teenage birth**
Tarapacá	0.038	-0.013
Antofagasta	0.038	-0.016
Atacama	0.042	-0.012
Coquimbo	0.040	-0.016
Valparaíso	0.042	-0.014
Libertador General Bernardo O'Higgins	0.028	-0.010
Maule	0.034	-0.012
Biobío	0.039	-0.013
Araucanía	0.044	-0.013
Los Lagos	0.035	-0.014
Aysén del General Carlos Ibáñez del Campo	0.019	-0.015
Magallanes y de la Antártica Chilena	0.033	-0.009
Metropolitana de Santiago	0.043	-0.015
Los Ríos	0.039	-0.013
Arica y Parinacota	0.026	-0.012
Ñuble	0.037	-0.008

Notes: The table reports the association between an extra year of parents' schooling and the likelihood of completing at least one year of tertiary education, as well as the likelihood of having a child as teenager for females (computed using an OLS regression). The samples include individuals aged between 19 and 21 (left) and 15 and 19 (right). Rows are sorted by the official designated number that each region used to have until 2018.

The selection of correlates follows considerations of data availability and is based on previous research, particularly Alesina et al. (2021), Chetty et al. (2014) and Van der Weide et al. (2024). An important caveat is that this analysis should not be interpreted as causal. The sole purpose is to document stylized facts that can later be used to theoretically model or estimate empirically the mechanisms behind local differences in intergenerational mobility.

4.1 Bivariate Associations

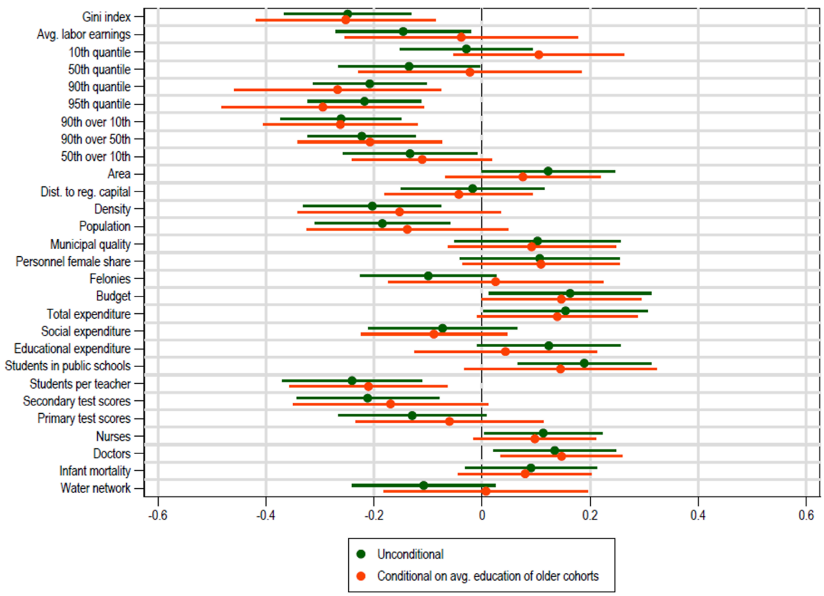
This section explores the bivariate relationship between indicators of IGM and a broad set of variables, including income inequality, governance, public investment, education, and health services, among others. The focus is on relative mobility measured as one minus the regression coefficient for simplicity and because it is arguably the most widely used indicator. However, a summary table with results for the remaining indicators can be found in the Appendix (Table A5). The definition of the correlates and their data sources are listed in the Appendix (see Table A1). To allow for a lag between the contextual environment and the observed level of mobility, all variables are measured in the year 2010.[25]

Figure 4 reports the coefficients from regressions with relative mobility (standardized to have mean 0 and variance 1) as the dependent variable and a given correlate (standardized to have mean 0 and variance 1) as the independent variable, which are labeled as unconditional, and then the same regression controlling for average education of the older cohorts.[26] This analysis provides an initial understanding of how these factors relate to IGM before applying a selection method based on their predictive ability. The unconditional estimates provide a descriptive perspective on the raw correlations between mobility and each variable, while the conditional estimates help isolate the effect above and beyond the average educational attainment of parents.

[25] The exception in terms of year of measurement is population, which is computed using Census 2017.

[26] Following the approach used in Alesina et al. (2021).

FIGURE 4
CORRELATES OF RELATIVE MOBILITY $(1-\beta)$
AT THE MUNICIPALITY-LEVEL



Notes: The figure reports the coefficients from regressions with relative mobility at the municipality level measured as one minus the regression coefficient (standardized to have mean 0 and variance 1) as the dependent variable and a given correlate (standardized to have mean 0 and variance 1) as the independent variable (in green, labeled as unconditional), and the coefficients from the same regression controlling for average education of the older cohorts (in orange). 95% confidence intervals are included.

Income distribution. Higher levels of inequality are often associated with lower mobility, as economic advantages and disadvantages persist across generations (Corak 2013). I include indicators such as the Gini index, specific income quantiles (10th, 50th, 90th, 95th), and income ratios (90/10, 90/50, 50/10) to capture different dimensions of inequality. The hypothesis is that greater income disparity, particularly between the richest and poorest households, limits access to high-quality education and economic opportunities, thereby reducing mobility.

I find that the Gini index, 90th quantile, 95th quantile, 90/10 ratio, and 90/50 ratio are all negatively and significantly correlated with relative mobility at the 5% level. These results suggest that in Chile, higher intergenerational mobility in education is more strongly associated with lower levels of income inequality in the upper half of the income distribution. This contrasts with findings from Corak (2020), which show that in Canada, mobility in income is more associated with inequality in the lower half of the income distribution. However, these results align with country-level evidence reported by Narayan et al. (2018), showing that income inequality is positively associated with intergenerational persistence in education, meaning it is negatively associated with relative mobility. These findings suggest that the well-documented relationship between inequality and mobility at the international level may also hold within countries (see Figure A13 in the Appendix).[27]

To complement this analysis, I also examine the relationship between relative mobility and inequality in education among parents (individuals aged 40–60 at the time of the Census). This exercise is equivalent to constructing a "Great Gatsby curve" for education within Chile, and I find a negative relationship between educational inequality and mobility. This suggests that the relationship documented across countries by the Narayan et al. (2018) also holds within countries (see Figure A14 in the Appendix).

Demographic and geographic characteristics. Spatial and demographic factors can influence mobility by affecting access to education, labor markets, and public services. I consider variables such as population size, density, municipality area, and distance to the regional capital. The hypothesis is that larger and more urbanized municipalities, or those closer to economic centers, may offer better educational resources and job opportunities, leading to higher mobility. However, none of these demographic variables exhibit a statistically significant relationship with mobility. This suggests that within-country variations in population distribution and geography do not play as strong a role in educational mobility.

Institutional and governance quality. The quality of local governance and institutional strength may affect mobility by influencing the efficiency of public service delivery. I include indicators such as municipal governance quality, female representation in the public sector, and felonies. The hypothesis is that better governance and lower crime levels may create an environment more conducive to higher intergenerational mobility. I do not find a significant relationship between these governance indicators and mobility. This may indicate that variations in local government performance across Chile are not substantial enough to drive differences in mobility outcomes.

Public spending and social infrastructure. Public investment in education and social services can mitigate economic disadvantages and enhance mobility. I analyze municipal budgets, total public expenditure, social expendi-

---

27 It is worth noting that the administrative dataset used to construct measures of income inequality, which comes from the unemployment insurance system, only considers the formal sector. This could lead to an underestimation of income inequality, as informal workers are excluded.

ture, and educational expenditure. The hypothesis is that higher public spending, particularly on education, should be positively associated with mobility, as it can reduce barriers in access to schools.

None of the expenditure variables show a statistically significant correlation with mobility, which suggests that the overall level of public investment alone may not be sufficient to drive mobility. A possibility is that what matters more may be the efficiency and targeting of educational resources rather than total spending. Further analysis is needed to explore whether certain types of educational investment (e.g., early childhood programs, teacher training) are more effective in promoting mobility.

Education system characteristics. The education system plays a fundamental role in shaping intergenerational mobility, as access to quality schooling can mitigate socioeconomic disadvantages. I analyze indicators such as the share of students in public schools, student-to- and teacher ratios, and standardized test scores (primary and secondary). The hypothesis is that smaller student-to-teacher ratios and higher test scores should be positively associated with mobility, as they reflect better learning environments and stronger educational outcomes. Additionally, a higher share of students in public schools could be linked to either higher or lower mobility depending on the quality of public education relative to private alternatives. Consistent with the expectations, I find that the log of students per teacher is negatively and significantly correlated with mobility at the 5% level, suggesting that overcrowded classrooms hinder educational progression.[28] This result aligns with prior evidence showing that smaller class sizes improve student performance, particularly for disadvantaged groups.

Health and basic services. Health infrastructure and basic services can play a crucial role in early childhood development, which in turn influences long-term educational and economic outcomes. I examine indicators such as the number of doctors and nurses per capita, infant mortality rates, and access to water services. The hypothesis is that better healthcare access and lower infant mortality rates should be positively associated with mobility, as healthier early-life conditions contribute to better learning and cognitive development. Additionally, access to clean water may reduce health-related school absences, further supporting educational progress.

Consistent with the hypothesis, I find that the number of doctors per capita is positively and significantly correlated with mobility at the 5% level. This supports the view that access to healthcare contributes to better childhood development and long-term educational outcomes.[29]

---

28 Secondary test scores are also negatively correlated but marginally insignificant at the 5% when conditioning on average education of older cohorts.

29 Budget availability, total expenditure, and number of nurses are also positively correlated but only marginally insignificant at the 5% level.

4.2 Identifying Key Predictors Of Mobility Using Lasso

To identify the most relevant predictors of intergenerational mobility (measured as one minus the regression coefficient), we estimate a least absolute shrinkage and selection operator (LASSO) regression.[30] This regularization technique selects variables by shrinking some coefficients to zero while retaining those with the strongest predictive power. A key advantage of LASSO is its ability to handle high-dimensional data by filtering out less relevant predictors. However, an important consideration is that differences in measurement error across variables may influence the selection process. Some variables may be retained not only because of their predictive strength but also because they are measured with greater precision.

Figure 5 presents the full coefficient paths from the LASSO estimation, allowing the penalization parameter λ to range from 0 (corresponding to an OLS model where all variables are included) to infinity (where all coefficients shrink to zero). The optimal value of λ is indicated by the vertical red line, highlighting the set of correlates that remain nonzero after regularization. The results show that the strongest predictors of relative mobility include the 90/10 earnings ratio, municipality area, population density, municipal quality, personnel's female share, municipal budget, share of students in public schools, student-to-teacher ratio, primary test scores, number of nurses, number of doctors, and water network coverage.

Among these, the four most influential predictors are the share of students enrolled in public schools, municipal budget, population density, and student-to-teacher ratio. The first two variables exhibit positive coefficients, suggesting that a higher proportion of students in public schools and greater municipal spending are associated with higher mobility. In contrast, population density and student-to-teacher ratios display negative coefficients, indicating that mobility tends to be lower in more densely populated areas and in municipalities where classrooms are more crowded. These findings reinforce the role of education system characteristics and local government resources in shaping mobility outcomes.

Compared to the bivariate analysis in Figure 4, several variables that are significantly associated with IGM in the bivariate analysis, such as some of the income inequality indicators, are not kept based on their predictive power, except for the ratio 90/10. Conversely, other correlates such as municipal budget and school enrollment composition, which were not among the strongest correlates in the bivariate analysis, emerge as important predictors once accounting for interactions between variables.

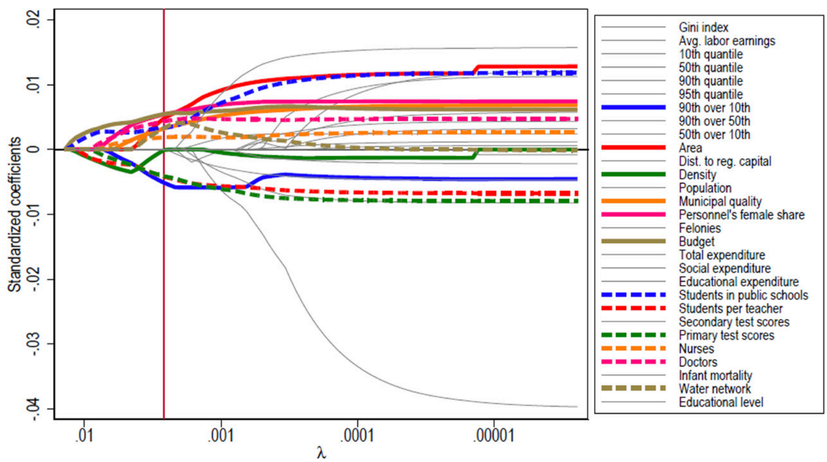
---

30 Results for all the indicators are available in Table A6 of the Appendix.

Interestingly, the strongest predictors of intergenerational mobility identified in the LASSO analysis closely mirror the three core dimensions of the Human Development Index: income, health, and education. This alignment suggests that mobility outcomes may reflect broader patterns of human development at the local level. Future work could explore more explicitly the connections between intergenerational mobility and composite development measures (e.g., Durlauf et al. 2022).

FIGURE 5

COEFFICIENT PATHS FROM LASSO ESTIMATES FOR RELATIVE MOBILITY $(1-\beta)$ AT THE MUNICIPALITY-LEVEL



Notes: The figure presents the full coefficient paths from the LASSO estimation, allowing the penalization parameter λ to range from 0 (corresponding to an OLS model where all variables are included) to infinity (where all coefficients shrink to zero). The optimal value of λ is indicated by the vertical red line, highlighting the set of correlates that remain nonzero after regularization.

Overall, these results suggest an important role for education system characteristics, local government capacity, and urbanization dynamics in shaping intergenerational mobility in Chile. Moreover, when the bivariate analysis and the LASSO selection are taken together, income inequality, the number of students per teacher, and the number of doctors are the correlates that stand out in both exercises.

5. Final Remarks

In this paper, I examine intergenerational educational mobility in Chile using full-count census microdata from a cohort born in the 1990s. I investigate intergenerational mobility in education using 8 indicators that relate to different aspects of the association between children's and parents' education, as well as the association between parents' education and other outcomes such as teenage birth or the ability to attend tertiary education.

I generate estimates at the municipality, regional, and country level. I document important within-country heterogeneity in intergenerational mobility as well as the association of parents' education with other outcomes. Using the municipality level estimates, I show that IGM is correlated with (and can be predicted with) labor earnings inequality, the number of doctors in the municipality, and the students per teacher ratio.

Although each of the eight indicators captures a distinct facet of educational mobility, I find that they offer a largely consistent view of the patterns observed within Chile. Relative mobility estimates based on regression $(1-\beta)$ and correlation $(1-\rho)$ are positively associated across municipalities and identify similar spatial patterns. The consistency extends to absolute mobility (α , $\overline{y}^{\geq}$), which tends to be higher in the same areas where relative mobility is also elevated. The directional measures ($P_{1,5}$, $P_{1,1}$, $P_{5,5}$) complement this picture, with low upward mobility from the bottom and high persistence at the top confirming that educational advantages and disadvantages are transmitted across generations.

Taken together, the evidence points to a country with moderate to low intergenerational educational mobility, depending on the metric. Chile performs reasonably well on absolute mobility (more than two-thirds of the cohort surpass their parents' education) but performs less well in relative and directional mobility, especially when assessed by correlation or the probability of moving from the bottom to the top of the distribution. This asymmetry highlights a key challenge: general educational gains do not necessarily translate into equality of opportunity for all. Structural inequalities, such as socioeconomic stratification in the school system and disparities in public service provision, appear to limit upward mobility for the most disadvantaged.

These findings open avenues for future research. One direction could be to explore how intergenerational mobility patterns documented in this article vary if local labor markets or commuting zones are considered rather than municipalities. Similarly, it could evaluate whether local patterns in coresidence rates impact the geographical variation in IGM.

References

Abramitzky, R., Boustan, L., Jácome, E. & Pérez, S. (2021), 'Intergenerational Mobility of Immigrants in the United States over Two Centuries', American Economic Review 111(2), 580–608.

Ager, P. & Boustan, L. P. (2021), 'The Intergenerational Effects of a Large Wealth Shock', American Economic Review 111(11), 3767–3794.

Alesina, A., Hohmann, S., Michalopoulos, S. & Papaioannou, E. (2021), 'Intergenerational Mobility in Africa', Econometrica 89(1), 1–35.

Alesina, A., Hohmann, S., Michalopoulos, S. & Papaioannou, E. (2023), 'Religion and Educational Mobility in Africa', Nature 618(7963), 134–143.

Asher, S., Novosad, P. & Rafkin, C. (2024), 'Intergenerational Mobility in India: New Measures and Estimates Across Time and Social Groups', American Economic Journal: Applied Economics 16, 66–98.

Card, D., Domnisoru, C. & Taylor, L. (2022), 'The Intergenerational Transmission of Human Capital: Evidence from the Golden Age of Upward Mobility', Journal of Labor Economics 40(S1), S39–S95.

Celhay, P. & Gallegos, S. (2015), 'Persistence in the Transmission of Education: Evidence across Three Generations for Chile', Journal of Human Development and Capabilities: A Multi-Disciplinary Journal for People-Centered Development 16(3).

Celhay, P. & Gallegos, S. (2025), 'Schooling mobility across three generations in six Latin American countries', Journal of Population Economics 38, 23.

Celhay, P., Sanhueza, C. & Zubizarreta, J. (2010), 'Intergenerational Mobility of Income and Schooling: Chile 1996-2006', Revista de Analisis Economico 25(2), 43–63.

Chetty, R., Hendren, N., Kline, P. & Saez, E. (2014), 'Where is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States', The Quarterly Journal of Economics 129(4), 1553–1623.

Corak, M. (2013), 'Income Inequality, Equality of Opportunity, and Intergenerational Mobility', Journal of Economic Perspectives 27(3), 79–102.

Corak, M. (2020), 'The Canadian Geography of Intergenerational Income Mobility', The Economic Journal 130(631), 2134–2174.

Cortés Orihuela, J., Díaz, J. D., Gutiérrez Cubillos, P., Montecinos, A., Troncoso, P. A. & Villarroel, G. I. (2023), 'Decentralizing the Chilean Miracle: Regional Intergenerational Mobility in a Developing Country', Regional Studies 57(5), 785–799.

Derenoncourt, E. (2022), 'Can You Move to Opportunity? Evidence from the Great Migration', American Economic Review 112(2), 369–408.

Deutscher, N. & Mazumder, B. (2023), ‘Measuring intergenerational income mobility: A synthesis of approaches’, Journal of Economic Literature 61(3), 988–1036.

Dodin, M., Findeisen, S., Henkel, L., Sachs, D. & Schüle, P. (2024), ‘Social Mobility in Germany’, Journal of Public Economics 232, 105074.

Durlauf, S. N., Kourtellos, A. & Tan, C. M. (2022), ‘The great gatsby curve’, Annual Review of Economics 14(1), 571–605.

Emran, M. S., Greene, W. & Shilpi, F. (2018), ‘When Measure Matters: Coresidency, Truncation Bias, and Intergenerational Mobility in Developing Countries’, Journal of Human Resources 53(3), 579–607.

Feigenbaum, J. J. (2018), ‘Multiple Measures of Historical Intergenerational Mobility: Iowa 1915 to 1940’, Economic Journal 128(612), F446–F481.

Fletcher, J. M. & Wolfe, B. L. (2009), ‘Education and labor market consequences of teenage childbearing: evidence using the timing of pregnancy outcomes and community fixed effects’, Journal of Human Resources 44(2), 303–325.

Francesconi, M. & Nicoletti, C. (2006), ‘Intergenerational Mobility and Sample Election in Short Panels’, Journal of Applied Econometrics 21, 1265–1293.

Gutiérrez, G. & Carrasco, A. (2021), ‘Chile’s enduring educational segregation: A trend unchanged by different cycles of reform’, British Educational Research Journal 47(6), 1611– 1634.

Hertz, T., Jayasundera, T., Piraino, P., Selcuk, S., Smith, N. & Verashchagina, A. (2007), ‘The Inheritance of Educational Inequality: International Comparisons and Fifty-Year Trends’, The B.E. Journal of Economic Analysis Policy 7(2).

Hilger, N. (2016), ‘The Great Escape: Intergenerational Mobility in the United States Since 1940’, NBER Working Paper (21217).

Mazumder, B. (2016), Estimating the intergenerational elasticity and rank association in the United States: Overcoming the current limitations of tax data, in ‘Inequality: Causes and consequences’, Vol. 43, Emerald group publishing limited, pp. 83–129.

Mizala, A. & Torche, F. (2012), ‘Bringing the schools back in: the stratification of educational achievement in the chilean voucher system’, International Journal of Educational Development 32(1), 132–144.

Muñoz, E. (2024), ‘The Geography of Intergenerational Mobility in Latin America and the Caribbean’, Economía LACEA Journal 23, 333–354.

Muñoz, E. & Siravegna, M. (2023), ‘When Measure Matters: Coresidence Bias and Intergenerational Mobility Revisited’, IDB Working Paper 01469.

Narayan, A., Van der Weide, R., Cojocaru, A., Lakner, C., Redaelli, S., Gerszon Mahler, D., Ramasubbaiah, R. & Thewissen, S. (2018), Fair Progress?:

Economic Mobility Across Generations Around the World, The World Bank.

Neidhöfer, G., Serrano, J. & Gasparini, L. (2018), 'Educational Inequality and Intergenerational Mobility in Latin America: A New Database', Journal of Development Economics 134, 329–349.

Nunez, J. I. & Miranda, L. (2010), 'Intergenerational Income Mobility in a Less-Developed, High-Inequality Context: The Case of Chile', The B.E. Journal of Economic Analysis Policy 10(1).

Oreopoulos, P. & Petronijevic, U. (2013), 'Making college worth it: A review of research on the returns to higher education'.

Sapelli, C. (2016), Cap 3: La Movilidad Intergeneracional de la Educacion en Chile, in 'Chile: ¿Más Equitativo? Una mirada a la dinamica social del Chile de ayer, hoy y mañana'.

Torche, F. (2005), 'Unequal But Fluid: Social Mobility in Chile in Comparative Perspective', American Sociological Review 70, 422–450.

Torche, F. (2021), Educational Mobility in the Developing World, in V. Iversen, A. Krishna & K. Sen, eds, 'Social Mobility in Developing Countries: Concepts, Methods, and Determinants', Oxford University Press.

Van der Weide, R., Lakner, C., Gerszon Mahler, D., Narayan, A. & Ramasubbaiah, R. (2024), 'Intergenerational Mobility around the World: A New Database', Journal of Development Economics 166, 103167.

APPENDICES

The appendix provides additional tables and figures, and other relevant information.

Table A1 lists the set of correlates that I use together with a short description and data sources.

Table A2 reports all the indicators computed by sub-populations (male vs. female, indigenous vs. non-indigenous, and urban vs. rural).

Table A3 reports some descriptive statistics of the estimates of IGM at the level of municipality.

Table A4 reports region-level estimates of rags to riches, intergenerational low, and intergenerational high comparing measures that assign individuals into quintiles using the distribution of educational attainment at the country level with measures that use the distribution of each region (those with the superscript "local").

Table A5 reports the statistical significance of the association between indicators of IGM and different correlates.

Table A6 presents the variables selected with a LASSO estimation using the optimal value of λ, highlighting the set of correlates that remain nonzero after regularization for each indicator of intergenerational mobility.

Figure A1 plots different measures of intergenerational mobility in education at country-level highlighting where Chile falls relative to Latin America and the Caribbean and the world.

Figure A2 plots different measures of intergenerational mobility in education for Chile com- pared to simple averages of regions for five different cohorts.

Figure A3 displays an histogram with the distributions of educational attainment of parents and children.

Figure A4 displays the evolution of mobility across birth cohorts in recent literature versus my estimate.

Figure A5 plots the average coresidence rate against IGM indicators at the municipality level.

Figure A6 plots the average coresidence rate by level of education.

Figure A7 plots the average coresidence rate by age.

Figure A8 displays the transition probabilities between educational attainment of parents and children (classified into three categories).

Figure A9 shows the cumulative distribution of the sample size by municipality.

Figure A10 displays the distribution of all the measures at municipality-level.

Figure A11 displays scatter plots comparing indicators of mobility (that use quintiles) using country level distribution of educational attainment vs. local distribution.

Figure A12 reports the results of the correlations with a set of variables using all the measures of IGM.

Figure A13 shows a binscatter plot between relative mobility and income inequality measured with the Gini coefficient at the municipality level.

Figure A14 shows a binscatter plot between relative mobility in education and educational inequality measured with the standard deviation of years of schooling at the municipality level.

Figure A15 maps the level of educational intergenerational mobility at the municipality level separating the country into north, metropolitan region, and south.

TABLE A1
COVARIATES

Label	Source	Description
Gini Index	UID	Gini Index
Average earnings	UID	Average earnings in the formal sector
10th quantile	UID	10th percentile of earnings in the formal sector
50th quantile	UID	50th percentile of earnings in the formal sector
90th quantile	UID	90th percentile of earnings in the formal sector
95th quantile	UID	95th percentile of earnings in the formal sector
Ratio 90-10	UID	Ratio 90th to 10th percentile of earnings in the formal sector
Ratio 90-50	UID	Ratio 90th to 50th percentile of earnings in the formal sector
Ratio 50-10	UID	Ratio 50th to 10th percentile of earnings in the formal sector
Area	SINIM	Log of the total surface of municipality
Distance to regional capital	SINIM	Log of the distance between the municipality and the regional capital
Population density per km2	SINIM	Log of population density per km2 by municipality
Population	SINIM	Log of municipality's estimated population in June 2012
Municipal professionalization	SINIM	Share of college educated workers in the municipality
Female Share in Municipality	SINIM	Share of female workers over the total workers in personnel of the municipality
Crimes	CEAD	Log of the number of crimes with greater social connotation
Budget availability	SINIM	Log of municipality's budget availability per capita
Total expenditure	SINIM	Log of municipality's total expenditure per capita
Social expenditure	SINIM	Log of the municipality's total expenditure in the social programs area per capita
Education expenditure	SINIM	Log of the municipality's total expenditure education programs
Students in public schools	ACE	Number of students enrolled in public schools over total enrollment
Students per teacher	SINIM	Log of students per teacher ratio in the municipal education system
Standarized test - secondary	ACE	Average score between math and language in SIMCE taken in high school
Standarized test - primary	ACE	Average score between math and language in SIMCE taken in 4th grade
Nurses by 100K inhabitants	SINIM	Log of number of nurses by 100.000 inhabitants within the municipality
Doctors by 100K inhabitants	SINIM	Log of number of doctors by 100.000 inhabitants within the municipality
Infant mortality rate	SINIM	Number of children under 1 year of age who die for every 1.000 live births
Water network	SINIM	Percentage of homes connected to drinking water network in the municipality
Parental education	Census	Average education of individual older than 24 but younger than 66

Unemployment insurance database (UID) can be accessed at: https://www.spensiones.cl/apps/bdp/index.php.

National system of municipal information (SINIM) can be accessed at: http://datos.sinim.gov.cl/datos municipales.php.

Center for crime studies and analysis (CEAD) can be accessed at: http://cead.spd.gov.cl/estadisticas-delictuales/.

Research unit, education quality agency data (ACE) can be accessed at: https://informacionestadistica.agenciaeducacion.cl/#/bases.

Census 2017 data can be requested from the National Institute of Statistics at: https://www.ine.cl.

TABLE A2
IGM AT COUNTRY-LEVEL FOR SUBGROUPS

	Male	**Female**	**Non-indigenous**	**Indigenous**	**Urban**	**Rural**
α	9.049	10.129	9.535	9.881	9.622	9.476
β	0.688	0.742	0.710	0.748	0.717	0.718
$\bar{Y}$	11.126	11.123	11.260	10.247	11.336	9.203
$\overline{y^{\geq}}$	0.628	0.707	0.663	0.690	0.657	0.754
ρ	0.624	0.658	0.640	0.676	0.652	0.649
$P_{1,5}$	0.068	0.108	0.089	0.082	0.092	0.071
$P_{1,1}$	0.419	0.310	0.365	0.367	0.359	0.390
$P_{5,5}$	0.331	0.378	0.357	0.306	0.353	0.358

The table reports country-level estimates of absolute mobility, relative mobility $(1-\beta)$, average parents' education, share of children with higher education than parents, relative mobility $(1-\rho)$, rags to riches, intergenerational low, and intergenerational high, respectively, all computed by subgroup. A description of the measures can be found in Table 2.

TABLE A3

DESCRIPTIVE STATISTICS OF IGM AT MUNICIPALITY-LEVEL

	Mean	SD	Min	Max	N
α	9.79	0.66	7.16	11.73	330
$1-\beta$	0.74	0.05	0.54	0.97	330
$\bar{Y}$	10.00	1.19	6.13	14.50	330
$\overline{y^{\geq}}$	0.71	0.07	0.48	0.90	330
$1-\rho$	0.68	0.06	0.50	0.96	330
$P_{1,5}$	0.09	0.03	0.02	0.22	312
$P_{1,1}$	0.36	0.06	0.09	0.57	313
$P_{5,5}$	0.33	0.05	0.10	0.46	190

The table reports descriptive statistics of estimates of absolute mobility, relative mobility (1 – β), average parents' education, share of children with higher education than parents, relative mobility (1 – ρ), rags to riches, intergenerational low, and intergenerational high, respectively, all of them at the municipality-level. I omit estimates with less than 50 observations. A description of the measures can be found in Table 2.

TABLE A4

REGION-LEVEL ESTIMATES OF IGM STATISTICS

Region	$P_{1,5}$	$P_{1,5}^{local}$	$P_{1,1}$	$P_{1,1}^{local}$	$P_{5,5}$	$P_{5,5}^{local}$
Tarapacá	0.10	0.11	0.37	0.33	0.32	0.31
Antofagasta	0.08	0.10	0.40	0.36	0.31	0.33
Atacama	0.07	0.10	0.40	0.34	0.29	0.36
Coquimbo	0.07	0.08	0.38	0.34	0.32	0.31
Valparaíso	0.09	0.10	0.35	0.36	0.34	0.33
Libertador General Bernardo O`Higgins	0.10	0.10	0.34	0.34	0.31	0.31
Maule	0.09	0.09	0.36	0.37	0.32	0.31
Biobío	0.11	0.09	0.33	0.36	0.37	0.32
Araucanía	0.07	0.07	0.38	0.39	0.37	0.36
Los Lagos	0.07	0.08	0.41	0.38	0.31	0.32
Aysén del General Carlos Ibáñez del Campo	0.05	0.09	0.44	0.35	0.23	0.32
Magallanes y de la Antártica Chilena	0.10	0.10	0.30	0.31	0.30	0.33
Metropolitana de Santiago	0.09	0.09	0.37	0.36	0.36	0.36
Los Ríos	0.06	0.06	0.38	0.37	0.34	0.35
Arica y Parinacota	0.14	0.14	0.28	0.30	0.31	0.32
Ñuble	0.11	0.09	0.33	0.35	0.36	0.27

The table reports region-level estimates of rags to riches, intergenerational low, and intergenerational high (a description of the measures can be found in Table 2). It compares measures that assign individuals into quintiles using the distribution of educational attainment at the country level with measures that use the distribution of each region (those with the superscript “local”).

TABLE A5

CORRELATES OF INTERGENERATIONAL MOBILITY. ALL THE INDICATORS

	Relative mobility	Absolute mobility	Average education	Relative mobility	Above parents	Rags to riches	Intergen. low	Intergen. high
	$(1-\beta)$	(α)	$(\bar{Y})$	$(1-\rho)$	$(\overline{y^{\geq}})$	$(P_{1,5})$	$(P_{1,1})$	$(P_{5,5})$
Gini index	⊗	⊗	×	⊗	⊗	∘		
Avg. labor earnings	×		×		×	×		×
10th quantile			⊗		⊗	×		∘
50th quantile	×		×		×	×		×
90th quantile	⊗	⊗	×	⊗	⊗	⊗		×
95th quantile	⊗	⊗	×	⊗	⊗	⊗		×
90th over 10th	⊗	⊗	×	⊗	×	∘		×
90th over 50th	⊗	⊗	×	⊗	⊗			
50th over 10th	×		⊗		⊗	×		∘
Area			×			⊗	⊗	
Dist. to reg. capital			×		∘	⊗	⊗	
Density	×		⊗		×			
Population	×		⊗	×	×	∘		
Municipal quality (prof)								

Personnel female share						⊗		
Felonies			×		⊗	◦		
Budget	×		⊗		⊗	×	×	
Total expenditure	×		⊗		×			
Social expenditure			⊗		⊗			
Educational expenditure			×		⊗			
Students in public schools	×	⊗	×	×	⊗	◦		
Students per teacher	⊗	⊗	⊗	⊗	×	×		
Secondary test scores	×		×	×	×	⊗		⊗
Primary test scores			⊗		◦			×
Nurses	×		×	×				
Doctors	⊗	⊗		⊗		×		
Infant mortality								
Water network			⊗		⊗			×

Notes: The table reports the statistical significance (at the 5% level) of the coefficients from regressions with an indicator of intergenerational mobility at the municipality level (standardized to have mean 0 and variance 1) as the dependent variable and a given correlate (standardized to have mean 0 and variance 1) as the independent variable (labeled as unconditional model), and the coefficients from the same regression controlling for average education of the older cohorts (labeled as conditional model). × if statistically significant in unconditional model. ◦ if statistically significant in conditional model. ⊗ if statistically significant in both models.

TABLE A6

LASSO VARIABLE SELECTION ACROSS IGM INDICATORS

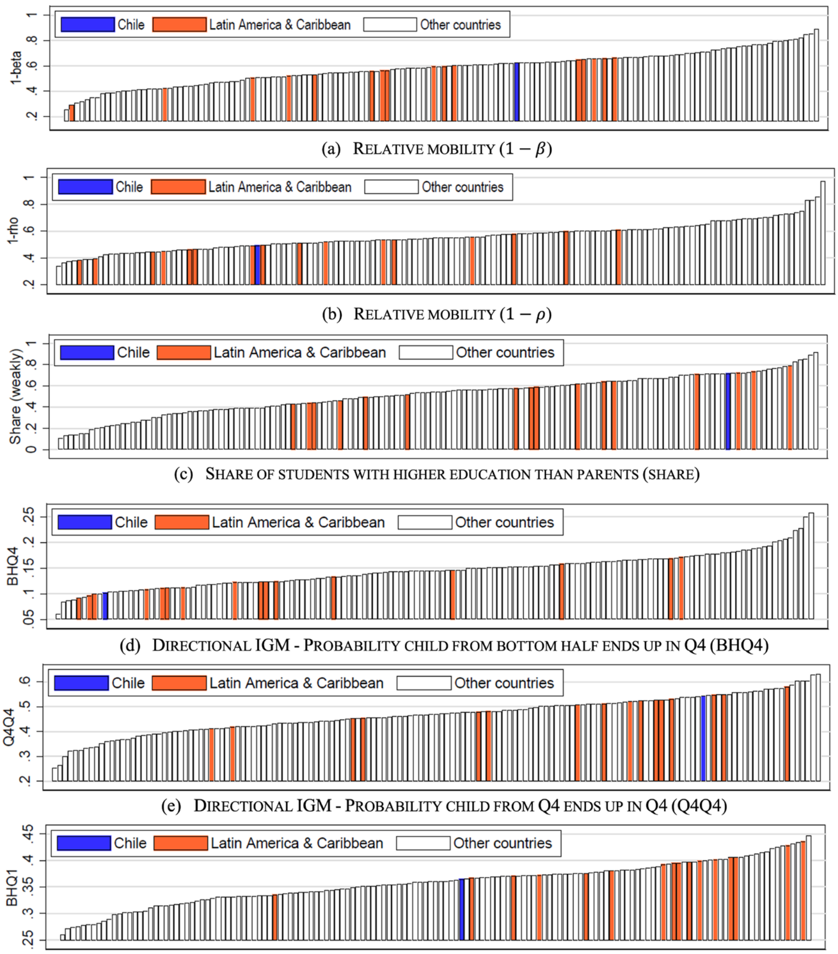
	Relative mobility	Absolute mobility	Average education	Relative mobility	Above parents	Rags to riches	Intergen. low	Intergen. high
	$(1-\beta)$	(α)	$(\bar{Y})$	$(1-\rho)$	$(\overline{y^{\geq}})$	$(P_{1,5})$	$(P_{1,1})$	$(P_{5,5})$
Gini index						×		
Avg. labor earnings								
10th quantile								
50th quantile					×	×		×
90th quantile								
95th quantile								
90th over 10th	×	×		×	×			
90th over 50th				×				
50th over 10th					×			
Area	×			×	×	×	×	×
Dist. to reg. capital		×			×	×		×
Density	×		×			×		×
Population			×					
Municipal quality (prof)	×	×		×	×	×		

Personnel female share	×	×		×	×	×	
Felonies					×	×	×
Budget	×	×		×		×	
Total expenditure							
Social expenditure				×	×	×	
Educational expenditure						×	
Students in public schools	×	×		×	×	×	
Students per teacher	×	×		×			
Secondary test scores					×	×	×
Primary test scores	×			×	×		
Nurses	×	×		×	×	×	
Doctors	×	×		×		×	
Infant mortality					×	×	
Water network	×	×	×	×	×		

Notes: The table presents the variables selected with a LASSO estimation using the optimal value of λ, highlighting the set of correlates that remain nonzero after regularization for each indicator of intergenerational mobility.

FIGURE A1

CHILE RELATIVE TO THE WORLD IN TERMS OF EDUCATIONAL IGM



(a) RELATIVE MOBILITY $(1-\beta)$

(b) RELATIVE MOBILITY $(1-\rho)$

(c) SHARE OF STUDENTS WITH HIGHER EDUCATION THAN PARENTS (SHARE)

(d) DIRECTIONAL IGM - PROBABILITY CHILD FROM BOTTOM HALF ENDS UP IN Q4 (BHQ4)

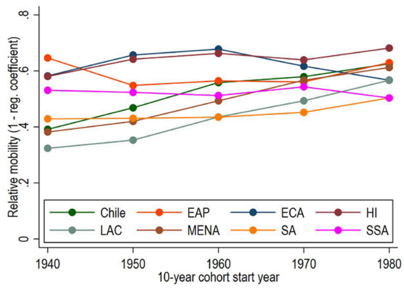
(e) DIRECTIONAL IGM - PROBABILITY CHILD FROM Q4 ENDS UP IN Q4 (Q4Q4)

(f) DIRECTIONAL IGM - PROBABILITY CHILD FROM BOTTOM HALF ENDS UP IN Q1 (BHQ1)

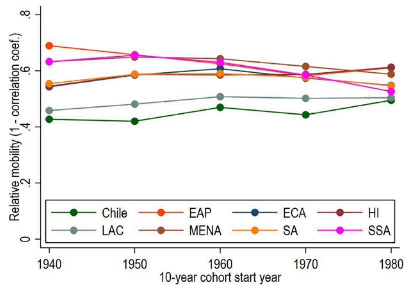
Source: Elaboration by the author with data from Narayan et al. (2018).

FIGURE A2

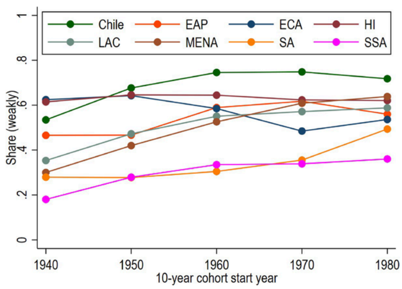
MOBILITY IN CHILE VERSUS AVERAGE BY REGION FOR FIVE COHORTS



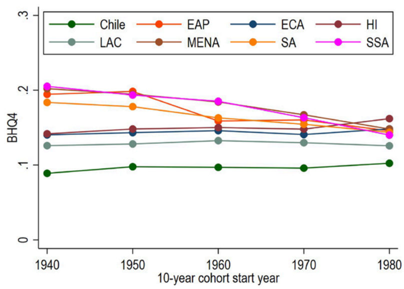
(a) RELATIVE MOBILITY $(1-\beta)$



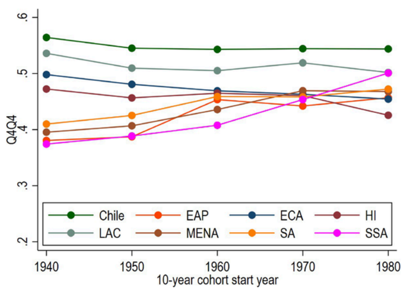
(B) RELATIVE MOBILITY $(1-\rho)$



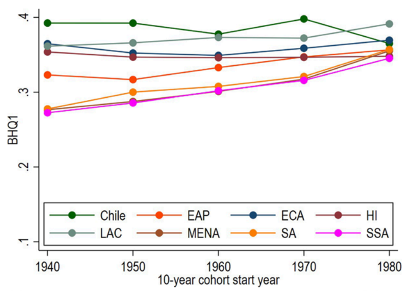
(C) SHARE WITH HIGHER EDUCATION THAN PARENTS



(D) PROB. CHILD FROM BOTTOM HALF ENDS UP IN Q4



(E) PROBABILITY CHILD FROM Q4 ENDS UP IN Q4



(F) PROB. CHILD FROM BOTTOM HALF ENDS UP IN Q1

Source: Elaboration by the author with data from Narayan et al. (2018). Regional averages are unweighted. Regions are EAP: East Asia & Pacific; ECA: Europe and Central Asia; HI: High income; LAC: Latin America and the Caribbean; MENA: Middle East and North Africa; SA: South Asia; SSA: Sub-saharian Africa.

FIGURE A3

HISTOGRAM OF EDUCATION

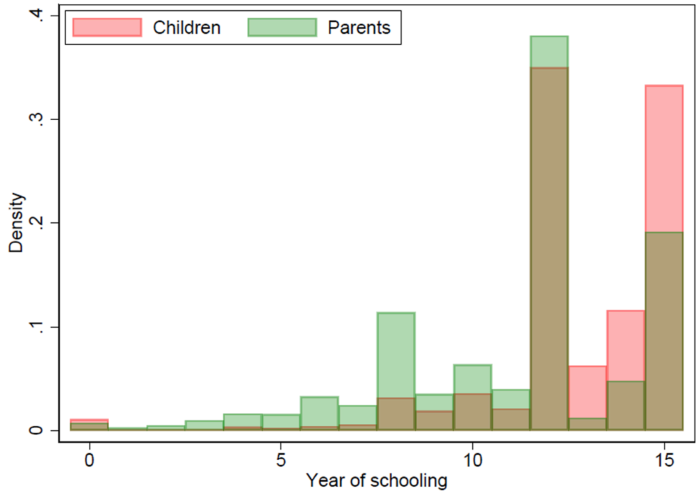
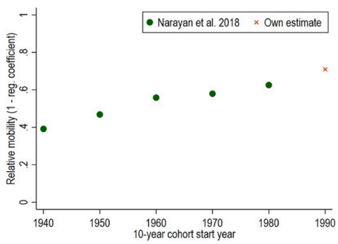
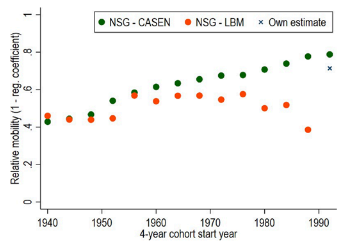


FIGURE A4

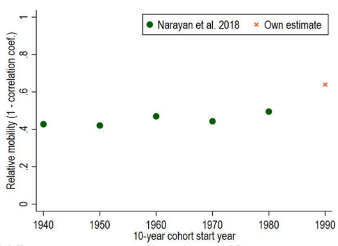
OWN ESTIMATES VERSUS RECENT LITERATURE AT THE COUNTRY LEVEL



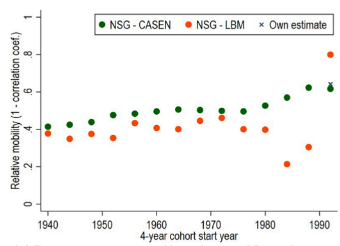
(A) RELATIVE MOBILITY $(1-\beta)$ FROM NARAYAN ET AL. (2018) AND OWN ESTIMATE.



(B) RELATIVE MOBILITY $(1-\beta)$ FROM NEIDHÖFER ET AL. (2018) AND OWN ESTIMATE.



(C) RELATIVE MOBILITY $(1-\rho)$ FROM NARAYAN ET AL. (2018) AND OWN ESTIMATE.



(D) RELATIVE MOBILITY $(1-\rho)$ FROM NEIDHÖFER ET AL. (2018) AND OWN ESTIMATE.

The figure shows estimates of intergenerational educational mobility obtained from regressing children years of schooling against parents' years of schooling, and the Pearson correlation coefficient between the same two variables. Narayan et al. (2018) uses CASEN survey while Neidhöfer et al. (2018) also uses Latinobarometro survey (LBM). The former uses 10-year cohorts, the latter uses 4-year cohorts (the most recent one is 1992-1995), and my estimate uses individuals approximately born between years 1991-1995. The last four cohorts using LBM survey contain smaller samples (831, 413, 179, and 24 observations), and hence are somewhat unreliable.

FIGURE A5

AVERAGE CORESIDENCE RATE VS. IGM INDICATORS AT THE MUNICIPALITY LEVEL

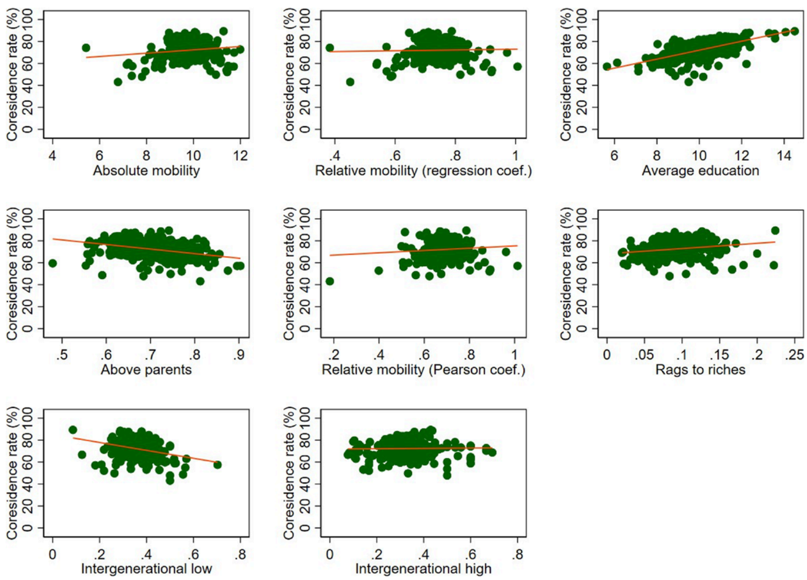


FIGURE A6

AVERAGE CORESIDENCE RATE BY LEVEL OF EDUCATION

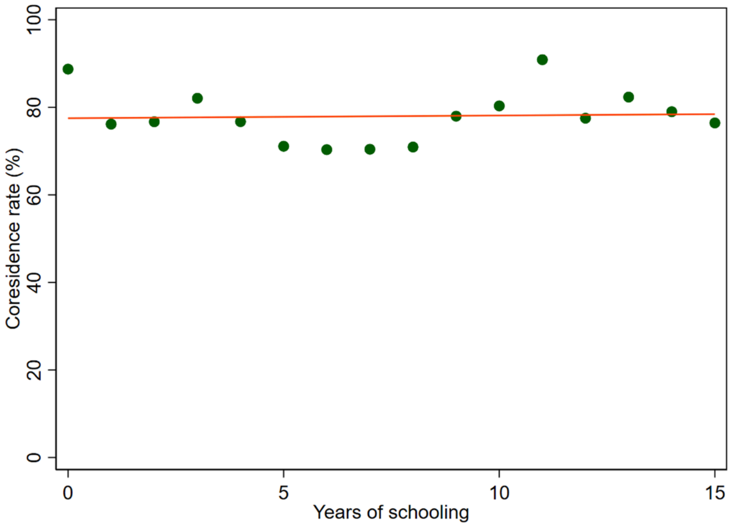


FIGURE A7

AVERAGE CORESIDENCE RATE BY AGE

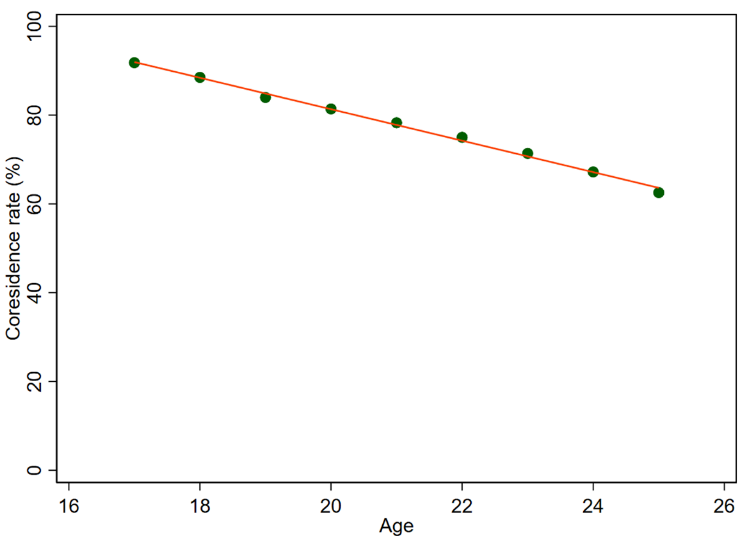


FIGURE A8

TRANSITION PROBABILITIES AT THE COUNTRY-LEVEL

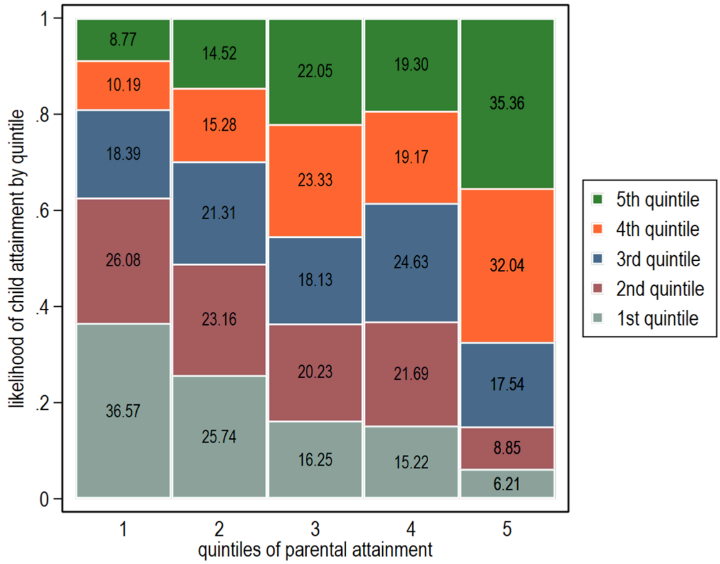


FIGURE A9

CUMULATIVE DISTRIBUTION OF THE SAMPLE SIZE

AT THE MUNICIPALITY LEVEL

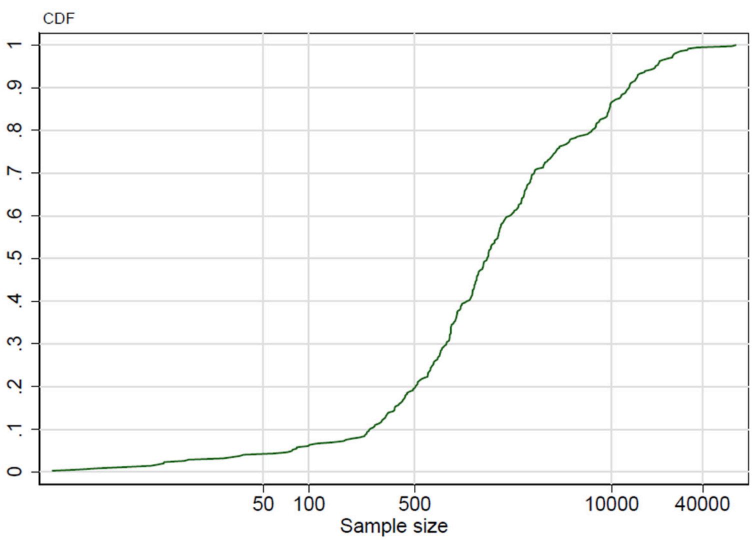
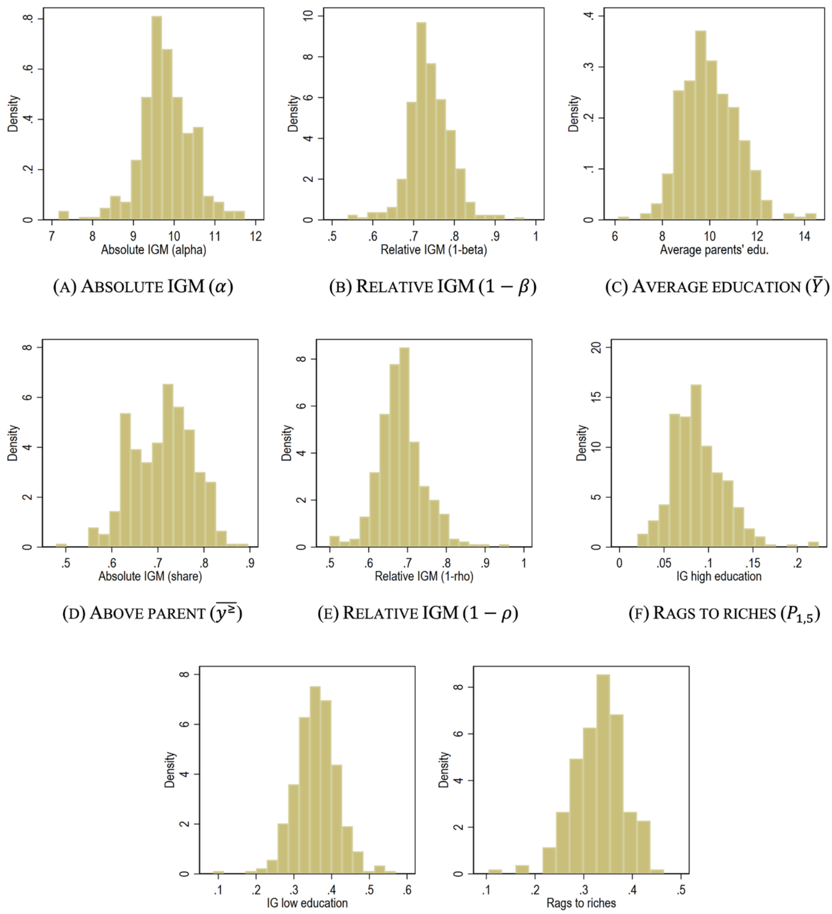


FIGURE A10

DISTRIBUTION OF MUNICIPALITY-LEVEL ESTIMATES



(A) ABSOLUTE IGM (α)　(B) RELATIVE IGM ($1-\beta$)　(C) AVERAGE EDUCATION ($\bar{Y}$)

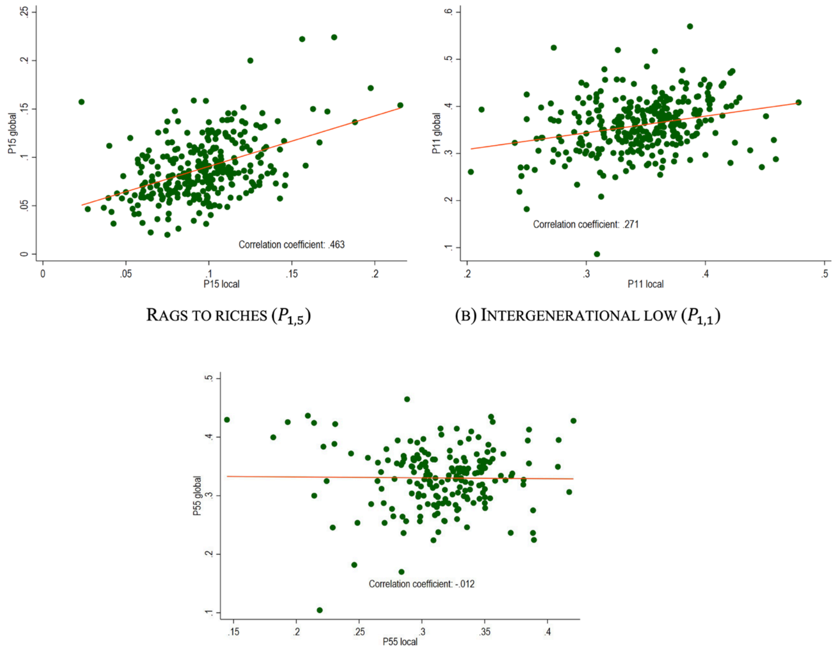
(D) ABOVE PARENT ($\overline{y^{\geq}}$)　(E) RELATIVE IGM ($1-\rho$)　(F) RAGS TO RICHES ($P_{1,5}$)

(G) INTERGENERATIONAL LOW ($P_{1,1}$)　(H) INTERGENERATIONAL HIGH ($P_{5,5}$)

These histograms show the distribution of the municipality-level estimates estimated with a sample of individuals of age 21-25 omitting municipalities with less than 50 individuals. For details about the indicators see Table 2.

FIGURE A11

COMPARISON OF INDICATORS USING COUNTRY LEVEL DISTRIBUTION OF EDUCATIONAL ATTAINMENT VS. LOCAL DISTRIBUTION



RAGS TO RICHES ($P_{1,5}$)

(B) INTERGENERATIONAL LOW ($P_{1,1}$)

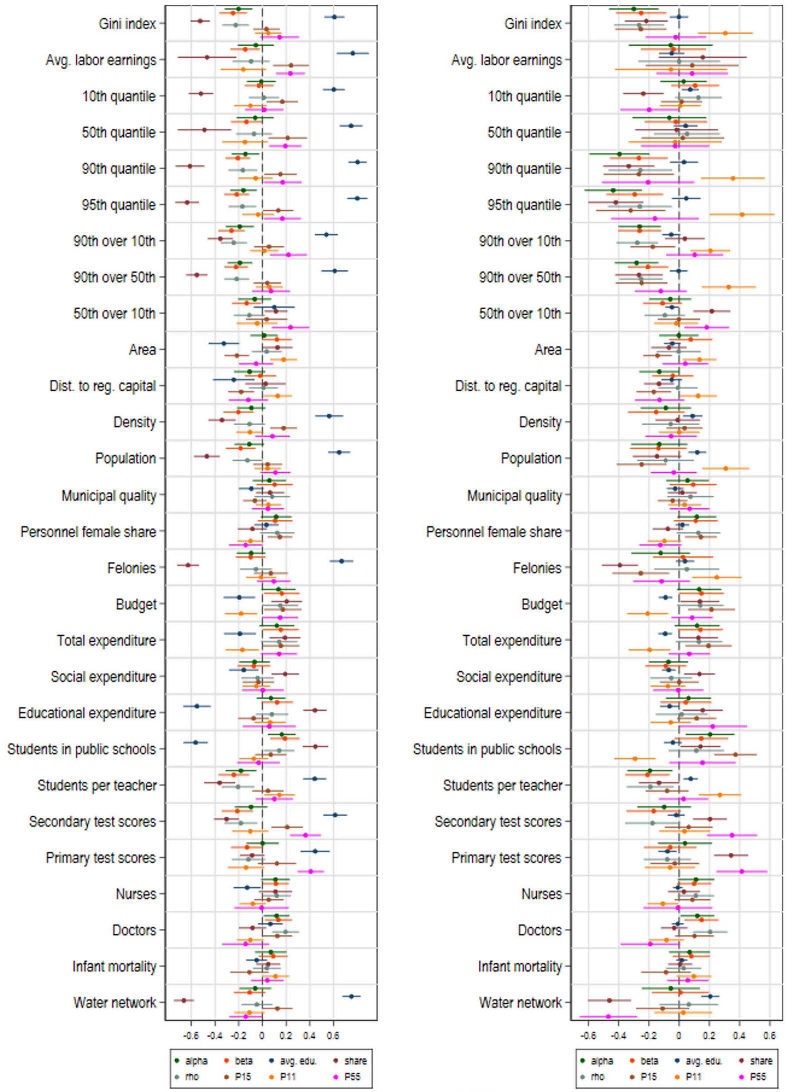
(C) INTERGENERATIONAL HIGH ($P_{5,5}$)

The figure compares estimates of rags to riches, intergenerational low, and intergenerational high measures computed using quintiles based on country-level educational attainment versus municipality-level attainment (denoted local). Each uses a sample of individuals of age 21-25 omitting municipalities with less than 50 individuals. For details about the indicators see Table 2.

FIGURE A12

CORRELATES OF THE IGM AT THE MUNICIPALITY-LEVEL

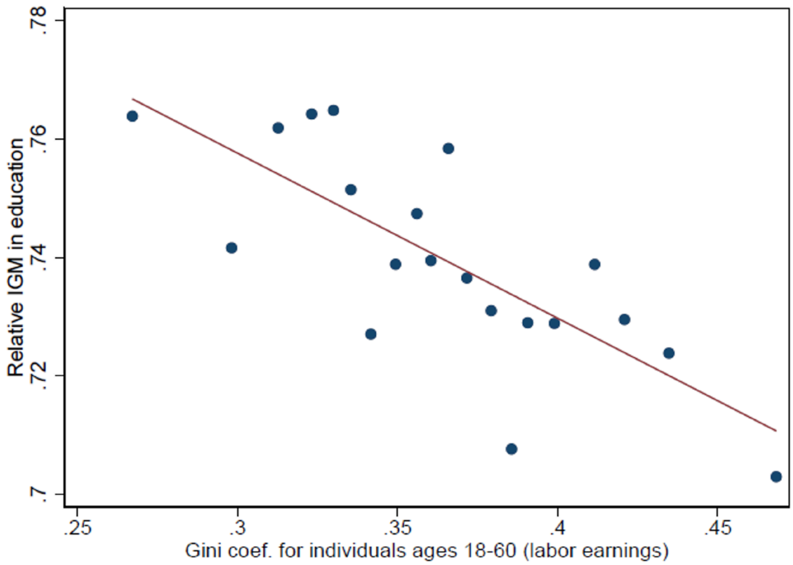
(ALL THE INDICATORS)



(A) UNCONDITIONAL (B) CONDITIONAL ON EDUCATION OF OLD COHORTS

FIGURE A13

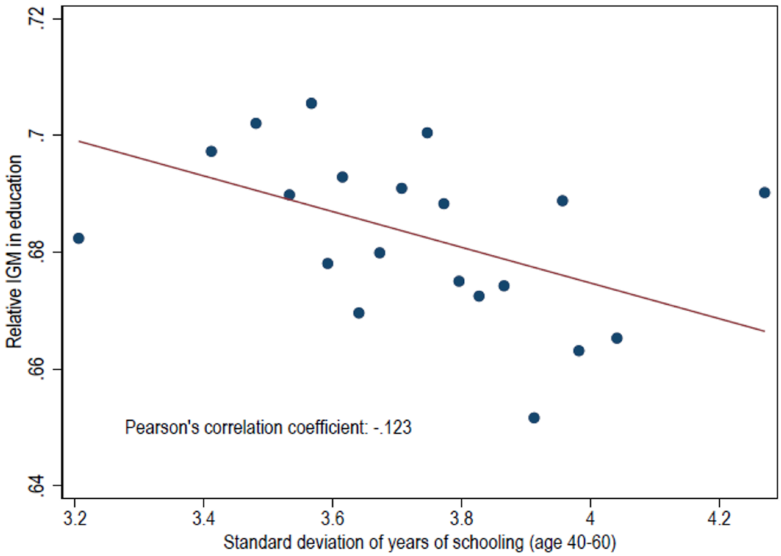
INTERGENERATIONAL MOBILITY IN EDUCATION VS. INCOME INEQUALITY



Notes: The figure shows a binscatter plot between relative IGM (measured as one minus the regression coefficient between child's years of schooling against parents' years of schooling) and the Gini coefficient computed using labor earnings in 2010 of individuals ages 18-60. Educational attainment is censored at 15 and the sample includes individuals with age between 21 and 25. Municipalities with less than 50 observations are not included.

FIGURE A14

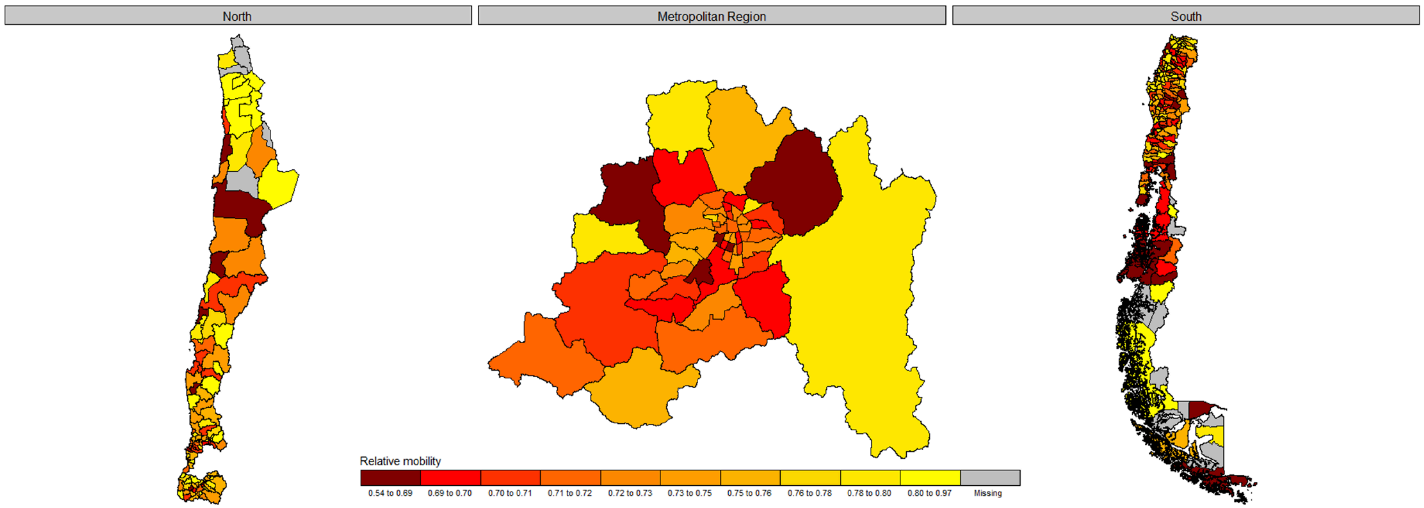
INTERGENERATIONAL MOBILITY IN EDUCATION VS. INCOME INEQUALITY



Notes: The figure shows a binscatter plot between relative IGM (measured as one minus the Pearson correlation coefficient between child's years of schooling against parents' years of schooling) and the standard deviation of years of schooling computed using individuals ages 40-60 that are used as parents. Educational attainment is censored at 15. Municipalities with less than 50 observations are not included.

FIGURE A15

INTERGENERATIONAL EDUCATIONAL MOBILITY WITHIN CHILE



(A) RELATIVE MOBILITY BY MUNICIPALITY - CHILE, 2017

Notes: The map plots relative IGM measured as one minus the regression coefficient (by municipality) between child's years of schooling (using age between 21 and 25) against parents' years of schooling. Educational attainment is censored at 15. Municipalities with less than 50 observations are left as missing.

Demand Elasticities for Selected Seasoning Commodities: An Almost Ideal Demand System with Instrumental Variables*

Elasticidades de la demanda para condimentos seleccionados: Un sistema de demanda casi ideal con variables instrumentales

MUHAMAD FATHUL MUIN **

Abstract

This study analyzes the consumption elasticities of five key seasoning commodities in Indonesia: cooking oil, red onion, garlic, red chili, and cayenne pepper. A Linear Approximate Almost Ideal Demand System (LA/AIDS) model is employed, incorporating instrumental variables to address potential endogeneity. The results indicate that the unconditional income and own-price elasticities are inelastic, with income elasticities ranging from 0.74 to 0.75 and own-price elasticities from –0.77 to –0.94. No significant evidence of substitution or complementarity among the seasonings is found. Furthermore, elasticity remain similar before and after the COVID-19 pandemic, and across regions with different economic sizes. However, regions known for spicy cuisines demonstrate higher elasticities than those with milder culinary traditions.

Keywords: *Seasoning commodities, Almost Ideal Demand System, Instrumental Variables, Indonesia.*

JEL Classification: *C36, D12, Q11.*

---

* I would like to express my gratitude to the three anonymous reviewers for their significant contribution to this paper. Their insightful comments and constructive suggestions further improved the quality of this work.

** Economics and Management School, Wuhan University, Hubei, China and China Institute of Development Strategy and Planning, Wuhan University, Hubei, China. Email: fathul.muin@whu.edu.cn / muhamad.fathul.muin@gmail.com. ORCID: https://orcid.org/0000-0002-7875-4537

DOI: http://dx.doi.org/10.4067/S0718-52862025000200309

Received: September, 2024 Accepted: May, 2025

Resumen

Este estudio analiza las elasticidades de consumo de cinco condimentos clave en Indonesia: aceite de cocina, cebolla morada, ajo, ají rojo y ají picante. Se emplea un modelo del Sistema de Demanda Casi Ideal en su versión Lineal Aproximada (LA/AIDS), que incorpora variables instrumentales para abordar posibles problemas de endogeneidad. Los resultados indican que las elasticidades no condicionadas respecto al ingreso y al precio propio son inelásticas, con elasticidades-ingreso que varían entre 0,74 y 0,75, y elasticidades-precio propias entre –0,77 y –0,94. No se encuentra evidencia significativa de sustitución o complementariedad entre los condimentos. Además, las estimaciones de elasticidad se mantienen relativamente estables antes y después de la pandemia de COVID-19, así como entre regiones con distintos tamaños económicos. Sin embargo, las regiones conocidas por sus cocinas picantes presentan elasticidades notablemente mayores que aquellas con tradiciones culinarias más suaves.

Palabras clave: *Condimentos, Sistema de demanda casi ideal, Variables instrumentales, Indonesia.*

Clasificación JEL: *C36, D12, Q11.*

1. Introduction

Studies on consumption patterns using demand systems have covered a wide range of topics. These include food consumption (Bakhtavoryan & Capps, 2024; Bilgic & Yen, 2013; Marioni et al., 2022; Roosen et al., 2022; Zhuang & Abbott, 2007), energy consumption (Burke & Abayasekara, 2018; Deryugina et al., 2020; Fouquet, 2014; Goetzke & Vance, 2021; Jin & Kim, 2022), health expenditure (Casabianca et al., 2022; Ellis et al., 2017; Farag et al., 2012; Z. Zhou et al., 2011), tourism expenditure (Fleissig, 2021; Fredman & Wikström, 2018; Gatt & Falzon, 2014; Untong et al., 2014), and transportation (Seya et al., 2024; Ventura et al., 2022; S. Wang & Noland, 2021; Wardman, 2024). Meta-analyses have comprehensively examined the elasticity of animal-derived food and energy consumption (Bouyssou et al., 2024; Labandeira et al., 2017). However, certain commodities such as seasonings remain understudied.

Seasonings—although not classified as major staple foods—play a crucial role in shaping cuisine flavours worldwide (Sproesser et al., 2022). These ingredients, which include various spices and flavourings, substantially contribute to the unique culinary identities. For instance, Indian cuisine is renowned for its extensive use of diverse spices and seasonings (Basak et al., 2023), while Chinese culinary emphasize achieving harmonious flavours through the combination of rice with specific aromatics, such as welsh onion, ginger, and garlic (Zhou et al., 2024). Similarly, Malaysian cuisine is characterized by the prominent use of spices, resulting in rich and complex flavour combinations (Abidin et al., 2020). These examples illustrate the importance of seasoning in culinary traditions, beyond the addition of flavour.

My study will examine the consumption patterns of seasoning commodities in Indonesia, a country with a rich culinary heritage. This gastronomic heritage is exemplified by the Mustikarasa Recipe Book, which document 1,600 traditional cuisines from various regions in Indonesia. These culinary preparations feature a wide variety of ingredients and flavours, including red onion, garlic, ginger root, turmeric, galangal, candlenut, lemongrass, red chili, and cayenne pepper (Wijaya et al., 2020). Therefore, this study provides insights into the seasoning consumption patterns within culinary diversity.

This study aims to calculate income and own-price elasticities, as well as identify substitution-complementarity relationships through cross-price elasticities of seasonings in Indonesia. This study contributes to the body of knowledge in several ways. First, it addresses the often-overlooked seasoning consumption patterns that are relevant to other countries with distinctive food consumption characteristics, such as India, Ethiopia, Thailand, and China (Helgy Library, 2023; Sherman & Billing, 1999). Second, this study adopts a multi-stage budgeting process to estimate long-run unconditional elasticity coefficients by assuming separable preferences and stable group price across utility levels (Ramírez, 2013). This approach is particularly unusual in Indonesian demand-system studies. Third, the study incorporated instrumental variables to address endogeneity issues in the model, a method rarely used in demand system modelling (Colen et al., 2018). This method is intended to promote the use of instrumental variables in future research on demand systems. Additionally, this study provides valuable policy insights into how rising prices affect public consumption patterns.

For the remainder, this paper is organized as follows. Section 2 explains the importance of studying seasoning demand elasticity and provides further background. Section 3 presents a brief literature review of the seasoning consumption elasticity. Section 4 introduces the demand system framework and the AIDS model. Section 5 describes the methods used in this study. Section 6 describes the descriptive statistics, model results, heterogeneity analysis, robustness checks, and simulation analysis. Section 7 discusses the research findings and their policy implications. Finally, Section 8 concludes the study and provides recommendations for future research.

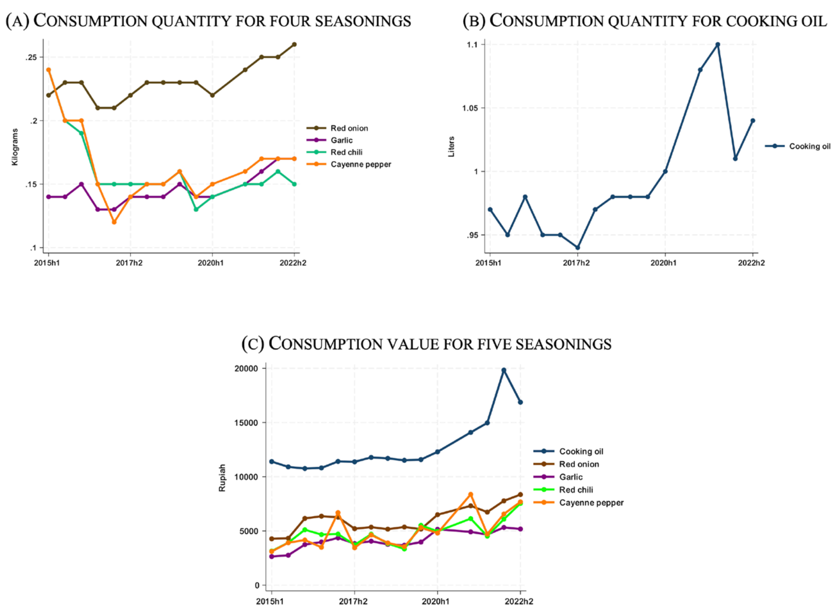
2. Seasonings And Indonesian Cuisines

The statistical report on household consumption identified at least 20 commonly consumed commodities. Seven of these were categorized as essential seasonings: cooking oil, red onions, garlic, red chili peppers, cayenne pepper, coconut, and sugar. Five of these seasonings are widely used in many Indonesian cuisines, such as in stir-fry, *pecel lele*, *rawon*, *nasi padang*, *coto*, *nasi liwet*, and many others. Cooking oil serves as a medium for sautéing basic seasonings and is used throughout the cuisine preparation. Red onions and garlic are often combined to enhance the overall cuisine flavour. Similarly, red chilies and cayenne peppers, although used differently, contribute heat and flavour to a range of dishes. Even, these chilies are also commonly associated with sambal—a traditional Indonesian condiment (Surya & Tedjakusuma, 2022).

Figure 1 shows the consumption patterns of the five major seasonings over eight years. Throughout the observation period, these commodities remained important to the Indonesian diet. Although the consumption quantities fluctuated over time, the overall trend showed a positive trajectory. Notably, red chilies and cayenne peppers deviated from this pattern, indicating a decline in the quantity consumed. However, when considering consumption value (quantity multiplied by price), all commodities—including red chili and cayenne peppers—showed an obvious upward trend.

FIGURE 1

CONSUMPTION OF FIVE SEASONINGS



Given their essential role in Indonesian cuisine, it is unsurprising that the government closely monitors the price fluctuations of these five seasonings. These seasonings—cooking oil, red onion, garlic, red chili, and cayenne pepper—contributed significantly to inflation in 2022. As shown in Table 1, these seasonings experienced multiple price increases throughout the year, with red onions having the highest frequency, with nine occurrences. Several factors contribute to these recurring price increases, including supply chain disruption and crop failure. These fluctuations significantly impact household budgets, prompting the government to oversee prices and ensure the affordability of these essential ingredients.

TABLE 1

PRICE MOVEMENTS DURING 2022

Commodities	**Jan**	**Feb**	**Mar**	**Apr**	**May**	**June**	**July**	**Aug**	**Sep**	**Oct**	**Nov**	**Dec**
Cooking oil	+	–	+	+	–	–	–	–	–	+	○	–
Red onion	+	+	+	○	+	+	+	–	–	+	+	+
Garlic	○	○	+	+	○	–	–	–	○	○	○	○
Red chili	–	–	+	–	–	+	+	–	–	+	–	+
Cayenne pepper	○	–	+	–	–	+	+	–	–	+	–	–

Source: BPS (2022, processing).

Source: sign + denote price increase.
sign – denotes price decrease.
sign ○ denotes constant price.

The impact of rising food prices is inseparably linked to a decline in purchasing power. When income levels remain constant, higher seasoning prices can lead to a reduction in the total quantity of goods purchased or to a shift towards the consumption of lower-quality goods. Consequently, rising prices are often associated with a decline in welfare level. However, there is a lack of data to accurately measure the impact of price increases on consumption patterns in Indonesia. The Socioeconomic Survey (SUSENAS) is the only source of comprehensive per capita consumption data covering expenditures on multiple commodities. However, this survey was conducted on a semiannual basis. Conversely, data for commodity prices are collected more frequently and comprehensively, with weekly collections and being aggregated into a monthly price index. This difference in the frequency of data collection poses significant challenges in establishing a causal relationship between price increases and changes in consumption patterns.

Despite data limitations, analyzing income and price elasticities of demand—both own-price and cross-price—for the five major seasonings remains crucial. In particular, understanding complementary goods through cross-price elasticity is important because these relationships can potentially reduce the purchasing power of other goods. From a public policy perspective, mapping complementary goods can facilitate the formulation of a more comprehensive strategy to mitigate food inflation. Therefore, this study is valuable for informing policy formulation and managing inflationary pressures on seasoning prices.

3. Previous Studies

The existing literature on seasoning consumption elasticity is relatively limited. A thorough review of the available research revealed that only four studies addressed this issue. Among these studies, two examined consumption elasticity in India, while others investigated the United States and Indonesia. This lack of research highlights the need further to explore seasoning consumption patterns across regions and markets.

Parappurathu and Mathur (2006) conducted a comprehensive study of spice demand in India—the country with the largest producer and consumer of spices. Using household data from nationwide surveys in 1987-1988 and 1999-2000, they estimated the demand elasticities. This study employed a multi-stage budgeting framework to model consumer behaviour and used a double-log regression model estimated through Ordinary Least Squares (OLS). The findings revealed that expenditure elasticities for spices in India were positive and inelastic. They also found that lower-income groups would increase their expenditure on spices more than higher-income groups in response to positive changes in income. Notably, the study found no significant difference in spice consumption between rural and urban households, suggesting that consumption patterns were relatively unchanged across regions.

Srivastava et al. (2013) conducted a similar study in India that examined the consumption patterns of six major spice commodities—dry chili, garlic, ginger, pepper, tamarind, and turmeric—using data from 2004 to 2005. The research employed Ordinary Least Squares (OLS) estimation and included demographic variables as controls, such as age, household size, proportion of children, and education. The results indicate that dry chili has the highest income elasticity among the spices studied, with demand increasing more rapidly as household income increases in both urban and rural areas. Conversely, ginger has the lowest income elasticity among the commodities examined.

In contrast to the two previous studies, Nguyen et al. (2019) used the Rotterdam model to assess the United States import demand for source-differentiated spices. Using monthly import data by region from January 1990 to April 2018, this study revealed that imported spices are normal goods in the U.S. market. Their study demonstrated a positive correlation between increases in real income and higher demand for spice imports, particularly those from Asia and South America. In terms of price sensitivity, U.S. demand for spices showed the highest elasticity to North American imports, while demand for Asian and South American spices was stable and less price-sensitive. The study further indicates that an increase in North American spice prices would likely result in a decrease in U.S. demand for these products and a subsequent shift to substitutes from alternative sources.

The most relevant study, both in terms of regional scope and the commodities examined, was conducted by Hamzah and Huang (2023). Using the Quadratic Almost Ideal Demand System (QUAIDS), they examined food consumption patterns in five regions using microdata from 2018. The study covered 12 major commodities in Indonesia, including rice, meat, eggs, red onions, garlic, chili, fish, cooking oil, white sugar, flour, processed foods, and miscellaneous food items. They implemented a two-step budgeting procedure that included demographic variables such as age, education level, household size, and rural-urban classification. The results indicated an absence of heterogeneity in the regional classification.

In summary, the first two studies primarily examined the consumption of seasonings specific to India such as dry chili, tamarind, and turmeric, which are not necessarily consumed in other regions. The third study used aggregate data without specifying the types of seasoning consumed, focusing instead on regional elasticity. The fourth study offered a more comprehensive approach by incorporating commodities of interest, including red onions, garlic, chilies, and cooking oil. However, these four studies have certain limitations that my study aims to address. The improvement encompasses the use of a three-stage budgeting mechanism, the application of instrumental variables, using macro household consumption data with a time series, and focusing on specific commodities with specific uses.

4. Demand System

4.1 Static Demand Analysis

Demand systems based on product space are frequently used in consumption studies. However, this methodological approach has several limitations (Pakes, 2021). First, it assumes that researchers possess the ability to accurately identify product choices from the perspective of the consumer. Second, it assumes that consumers have a comprehensive understanding of their products. Moreover, static demand systems require a significant number of coefficients—$2n^2+n$ for the n goods under study—and are incapable of forecasting the demand for new products. Based on these limitations, Pakes (2021) proposed a shift from product space to characteristic space analysis. This shift offers a more dynamic and flexible framework for understanding consumer behaviour and overcoming the limitations of traditional product space-based demand systems.

To address this criticism, the selection of commodities for this study was based on food consumption data from 2015 to 2022. These data revealed that the five chosen commodities consistently demonstrated high consumption levels and are widely recognized as essential components of daily dietary intake. These commodities fulfil distinct culinary functions and are intuitively understood by consumers without explicit instructions. By focusing on commodity-level data, this study encompasses various brands and variants within each category, despite the published data being in aggregate form. This approach ensures the capture of new products—whether brands or variants—within the same commodity category.

Furthermore, this study employs a multi-stage budgeting scheme to analyse the five main seasoning commodities. The analysis was conducted in three stages, generating a total of 75 coefficients: 10 unconditional coefficients in the first stage, 10 conditional coefficients in the second stage, and 55 conditional coefficients in the third stage. Despite this comprehensive calculation, the study primarily focuses on unconditional coefficients from the third stage.

Therefore, the final number of coefficients is reasonable and ensures an effective interpretation.

To accommodate the characteristic space-based analysis, this study incorporated heterogeneity analysis by classifying regions based on cuisine variations. Certain provinces are known for their spicy flavors, which are closely associated with the consumption of red chili, cayenne pepper, cooking oil, red onions, and garlic. Conversely, other regions are assumed to have relatively lower consumption of these five seasonings. By categorizing provinces according to their culinary characteristics, I was able to analyze and compare spatial variations in seasoning use and flavor preferences across regions.

Of course, implementing a static demand system yields long-run coefficients that capture the consumption response over an extended period of time. Conversely, dynamic demand systems must be implemented to obtain short-term coefficients. However, a dynamic approach was not included in this study due to the unavailability of the STATA commands. In addition, the semi-annual data structure used in this study makes the short-term approach infeasible. The effective implementation of dynamic analysis for food commodities requires a shorter data structure, such as weekly or monthly, for both consumption and price data. Despite these limitations, the static demand system approach offers significant long-term perspectives on the consumption patterns.

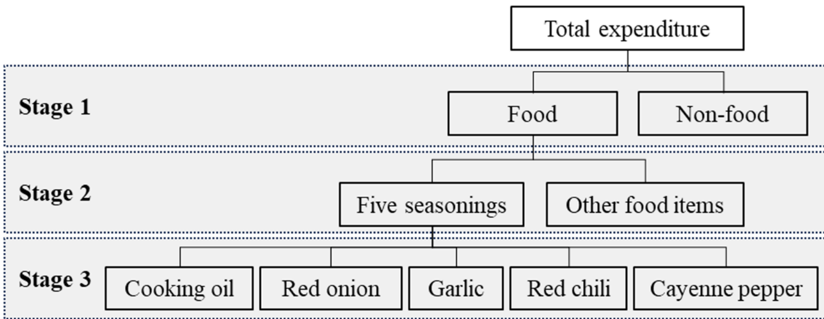
4.2 Multi-stage Demand System

Consumption allocation across income groups was extensively investigated by Chai et al. (2024) and Bonke & Browning (2009). However, the precise stages of household budgeting and expenditure allocation remain unclear. Various studies have proposed different approaches to analyse this process, including the adoption of a three-stage model (Dey et al., 2011; Edgerton, 1997; Ramírez, 2013; Rathnayaka et al., 2021; Wu et al., 2021) and a two-stage model (Hamzah & Huang, 2023; Hasiner & Yu, 2020; Y. Wang & Çakır, 2020; Widarjono & Rucbha, 2016). In addition, some researchers have focused on specific commodity groups without explicitly defining the stages (Čiderová & Ščasný, 2022; Díaz & Medlock, 2021; Suárez-Varela, 2020; Wongmonta, 2022).

As noted by Edgerton (1997), determining the stages of a consumption analysis remains subjective and questionable. Despite this challenge, multi-stage modelling offers significant advantages over single-stage demand analysis. Primarily, it allows for the calculation of unconditional elasticities, which more accurately reflect real-world consumer behaviour. In contrast, single-stage demand analysis, with its narrow focus on specific goods, yields conditional coefficients with limited practical utility. These conditional elasticities depend on higher-level demand functions, which preclude their direct application in understanding consumer behaviour. As a result, although multi-stage modelling is more complex, it provides a more accurate representation of actual consumption decisions.

In this study, a three-stage consumption model was employed, as illustrated in Figure 2. The first stage, grounded in Engel's law, examines the allocation of expenditure between food and non-food commodities. Engel's law, derived from the work of Ernst Engel, postulates that the share of income allocated to food decreases as income increases. Building on this foundation, then the second stage focuses on the allocation of food expenditures between the five seasonings and other food items. The final stage looks more closely at the allocation of expenditure among the five specific seasonings: cooking oil, red onion, garlic, red chili, and cayenne pepper. This hierarchical approach allows for a comprehensive analysis of consumer expenditure patterns, from broad categories to specific food items.

FIGURE 2
CONSUMPTION DECISION STAGE



4.2 Almost-Ideal Demand System (AIDS)

The Almost-Ideal Demand System (AIDS), introduced by Deaton and Muellbauer (1980), provides a framework for analysing consumer spending patterns. This model primarily requires expenditure and price data for its implementation. Overall, the estimation of the demand system equations involves three key variables: total consumption expenditure (M_{ht}), expenditure share (w_{iht}), and commodity prices (p_{iht}). In this study, I implemented the AIDS model for each stage by focusing on different commodity groups. Regarding scope, my study examined five seasoning commodities (*i=5*) across 33 provinces (*h=33*) in Indonesia, ranging from Aceh to Papua. For the data periods, this study encompasses 15 semi-annual periods from 2015h1 to 2022h2 (*t=15*).

To calculate the household share of commodity *i*, I used the following formula:

(1) $$w_{iht} = \frac{p_{iht} q_{iht}}{M_{ht}},\ 0 \leq w_{iht} \leq 1, \text{ and } \sum_{i=1}^{n} w_{iht} = 1$$

The AIDS formula used in this study was based on the works of Deaton and Muellbauer (1980) and Moschini (1995). Here, I adapted their approach to fit the specific data structure. Thus, the mathematical representation of this formula is as follows:

(2) $$w_{iht} = \alpha_i + \sum_{j=1}^{n} \gamma_{ij} \ln p_{jht} + \beta_i \ln\left(\frac{M_{ht}}{P_{ht}}\right) + \varepsilon_{iht}$$

(3) $$\text{with } \alpha_i = \alpha_{i0} + \alpha_{i1} t + \alpha_{i2} Z_{ht}$$

The model incorporates several key components and parameters to analyse household consumption patterns. At its core, the model uses ε_{iht} as an error term, P_{ht} denotes the translog price index, and Z_{ht} denotes the demographic variables, specifically total household members. The parameters in the model serve different functions: α_i represents the average share of good *i*, γ_{ij} measures the effect on the household share of good *i* of an increase in the relative price of good *j*, holding expenditure constant, and β_i indicates the effect on the household share of good *i* of an increase in real per capita expenditure. However, γ_{ij} and β_i are not direct measures of elasticity. Then to simplify the analysis, this study used the Linear Approximated AIDS (LA/AIDS) model. In this approach, the original translog price index P is replaced by the Stone index P, which is expressed by the following formula:

(4) $$\ln P_{ht} = \sum_{i=1}^{n} w_{it} \ln p_{iht}$$

Furthermore, demand theory requires that certain restrictions must be satisfied to ensure consistency with utility maximization.

(5) Additivity $$\sum_{i=1}^{n} \alpha_i = 1, \sum_{i=1}^{n} \beta_i = 0, \sum_{i=1}^{n} \gamma_{ij} = 0,$$
$$\sum_{i=1}^{n} \alpha_{i0} = 1, \sum_{i=1}^{n} \alpha_{i1} = 0, \sum_{i=1}^{n} \alpha_{i2} = 0$$

(6) Homogeneity $$\sum_{j=1}^{n} \gamma_{ij} = 0$$

(7) Slutsky symmetry $$\gamma_{ij} = \gamma_{ji}$$

From the above specification, the conditional expenditure elasticity (η_i), conditional uncompensated (Marshalian) price elasticity (e_{ij}^{u}), and conditional compensated (Hicksian) price elasticity (e_{ij}^{c}) in each level are formulated as follows:

(8) $$\eta_i = 1 + \left(\beta_i / w_i\right)$$

(9) $$e_{ij}^{u} = -\delta_{ij} + \left(\gamma_{ij} / w_i\right) - \beta_i \left(w_j / w_i\right)$$

(10) $$e_{ij}^{c} = -\delta_{ij} + \left(\gamma_{ij} / w_i\right) + w_j$$

(11) $$e_{ij}^{c} = e_{ij}^{u} + \eta_i w_j$$

where δ_{ij} is the Kronecker delta with a value equal to one when $i=j$ and zero when $i \neq j$. Then β_i and γ_{ij} are the estimated parameters.

As previously mentioned, this study proposes a three-stage consumption decision-making process that yields two types of elasticity: conditional and unconditional. Conditional elasticity is calculated at the second and third stage using equations 8-10, with interpretations influenced by stage-specific budget constraints. Conversely, unconditional elasticity is derived by modifying the formulation according to Edgerton (1997) and Ramírez (2013), which are expressed as follows:

(12) $$\eta_i^{T} = \eta_{i(stage\,3)} \times \eta_{five\,seasonings\,(stage\,2)} \times \eta_{food\,(stage\,1)}$$

(13) $$\begin{aligned} e_{ij}^{u(T)} &= e_{ij(stage\,3)}^{c} + \left(w_{j(stage\,3)} \times \eta_{i(stage\,3)} \times e_{five\,seasonings\,(stage\,2)}^{c}\right) \\ &+ \left(w_{j(stage\,3)} \times w_{five\,seasonings\,(stage\,2)} \times \eta_{i(stage\,3)} \times \eta_{five\,seasonings\,(stage\,2)} \times e_{food\,(stage\,1)}^{u}\right) \end{aligned}$$

(14) $$\begin{aligned} e_{ij}^{c(T)} &= e_{ij(stage\,3)}^{c} + \left(w_{j(stage\,3)} \times \eta_{i(stage\,3)} \times e_{five\,seasonings\,(stage\,2)}^{c}\right) \\ &+ \left(w_{j(stage\,3)} \times w_{five\,seasonings\,(stage\,2)} \times \eta_{i(stage\,3)} \times \eta_{five\,seasonings\,(stage\,2)} \times e_{food\,(stage\,1)}^{c}\right) \end{aligned}$$

where η_i^{T}, $e_{ij}^{u(T)}$, $e_{ij}^{c(T)}$ represent unconditional expenditure elasticity, unconditional uncompensated (Marshalian) price elasticity, and unconditional compensated (Hicksian) price elasticity, respectively.

5. Methods

5.1 Data and Model Estimation

This study examined 33 provinces in Indonesia from 2015 to 2022, excluding Kalimantan Utara and other newly established provinces. The data were primarily obtained from two types of surveys published by BPS-Statistics Indonesia. The first was a semiannual Socioeconomic Survey (SUSENAS) conducted in March and September that provided data on consumption. The second was a monthly consumer price survey conducted in multiple cities in each province, which provided information on commodity prices. Additional data were obtained from tables available online on the BPS-Statistics Indonesia website (www.bps.go.id).

For the model estimation, I employed the STATA command developed by Lecocq and Robin (2015). STATA offers a range of demand system modelling options, including Cobb-Douglas, basic linear expenditure, translog, generalized translog, almost ideal demand system (AIDS), generalized AIDS, quadratic AIDS, and generalized quadratic AIDS. While these models are predominantly static, recent developments have focused on dynamic demand system modelling, particularly for AIDS. Blanciforti and Green (1983) incorporated habit effects based on Pollak and Wales (1969) to develop a dynamic version of AIDS, which has been applied in subsequent research (Rathnayaka et al., 2019; Selvanathan et al., 2024). Nevertheless, endogeneity remains a significant concern among economic researchers. Therefore, this study employed a static LA/AIDS model and addressed the endogeneity issue in the analysis.

5.2 Handling Endogeneity

AIDS modelling faces endogeneity challenges arising from two primary sources. First, the total expenditure variable was constructed using budget shares as the dependent variable. Second, price variables may be determined simultaneously by supply-demand interactions. These factors are likely to result in a correlation between the error term (ε_{iht}) and both total expenditure (M_{ht}) and price (p_{jht}) variables. To address this issue and ensure robust results, I employ two instrumental variables (IVs) in the third stage.

The selection of instrumental variables for price and income was based on relevance and data availability. For the price instrument, I used data of quantity production for palm oil, red onions, garlic, red chili, and cayenne peppers. Ideally, the production quantity of cooking oil in each province would be more appropriate. However, given the data limitations, data productivity for palm oil was used here as a proxy. Similarly, data on stock quantity—for red onion, garlic, red chili, and cayenne pepper—is preferable rather than production data. This is primarily because stock levels are a more accurate representation of the supply side effects. Again, due to data limitations, production data were ultimately used here. For income, I then used GDP per capita as an instrumental variable. Furthermore, following the approach of Lecocq and Robin (2015), this study estimated the model parameters using the iterated linear least-squares method. This method successfully addresses potential endogeneity concerns while working within the constraints of the available data.

To evaluate the strength of the instrumental variables, there are several statistical measures that can be used. Following Stock & Watson (2020) and Sanderson et al. (2021), the F-statistic is commonly used as the rule of thumb.

An F-value greater than 10 indicates a strong instrumental variable, while a value less than 10 suggests a weak instrumental variable. However, Keane & Neal (2023) propose a higher standard for instrument strength in empirical practice. Therefore, in my analysis, I consider not only the F-statistic but also the adjusted R-squared and t-statistics for each type of instrumental variable to provide a more comprehensive evaluation.

Table 2 presents the validity test results for the six instrumental variables used in this study. Among these, four variables showed an F-statistic of 10 and above. However, only two variables—cooking oil and income—showed relatively high adjusted R-squared values. Further, the identification based on the t-statistic reveals that only palm oil production and GDP per capita are statistically significant instrumental variables. Therefore, I can conclude that out of the six instrumental variables used, two can be classified as strong instrumental variables while the remaining four can be considered as weak instrumental variables.

TABLE 2

CHECKING FOR INSTRUMENT STRENGTH

Endogenous variables	**F-stat**	**Adj. R-squared**	**Instrumental variables**	**t-stat**
Cooking oil	52.04***	0.452	Palm oil production	-3.81***
Red onion	10.61***	0.135	Red onion production	0.03
Garlic	8.89***	0.113	Garlic production	1.16
Red chili	10.00***	0.127	Curly chili production	0.28
Cayenne pepper	7.65***	0.097	Cayenne pepper production	-0.10
Income	37.08***	0.369	GDP per capita	5.21***

Source: Data processed.

Note: *, **, and *** refer to the significant level in 10%, 5%, and 1%, respectively.

6. Results

6.1 Descriptive Analysis

The data presented in Table 3 explain the consumption patterns of the five seasonings, focusing on the average expenditure, share of consumption, prices, and quantities consumed. Among these, cooking oil is the dominant commodity, accounting for about 39 percent of total consumption. Although the price of cooking oil is reported to be the lowest among other seasonings, the quantity consumed exceeds that of other seasonings by up to 4-7 times. As a result, the total value of consumption in monetary units remains high. By contrast, the remaining four seasonings demonstrate modest consumption shares, ranging from 12 to 20 percent.

TABLE 3
DESCRIPTIVE STATISTICS

Commodities	Expenditure (rupiah)	Share (percentage)	Price (rupiah)	Quantity
Cooking oil	13,449.320	38.871	14,585.660	0.998[a)]
	(3,791.632)	(6.553)	(2,528.881)	(0.156)
Red onion	6,751.234	19.195	32,631.940	0.233[b)]
	(2,111.960)	(2.920)	(18,421.76)	(0.060)
Garlic	4,366.323	12.556	30,981.170	0.142[b)]
	(1,767.403)	(3.882)	(9,289.332)	(0.043)
Red chili	5,250.842	13.244	38,913.840	0.163[b)]
	(5,645.274)	(11.101)	(15,310.520)	(0.158)
Cayenne pepper	5,775.006	16.134	49,216.110	0.165[b)]
	(3,254.201)	(6.644)	(23,111.740)	(0.078)

Source: Data processed.
Note: Quantity for each commodity defined as a) Liter, b) Kilogram.
(…) refers to standard deviation.

A comparison between similar commodities reveals a different pattern of consumption. The consumption of cayenne pepper was 1.2 times higher than that of red chili, while the consumption of red onion was 1.5 times higher than that of garlic. However, in terms of quantity, both types of chilies had relatively similar levels. In terms of price, red onion and garlic were relatively similar, whereas cayenne pepper was 20 percent more expensive than red chili. This price difference explains the higher consumption share of cayenne peppers compared to red chili, despite their similar quantitative consumption.

6.2 Elasticity Calculation

Table 4 presents the elasticity coefficients for each stage, including income elasticity and both Marshallian (uncompensated) and Hicksian (compensated) own-price elasticities. The analysis reveals the different responses of food and the five major consumption types of seasoning to changes in budget allocation. Food consumption is inelastic to budget allocation, indicating that an increase in budget allocation leads to a proportionally smaller increase in food consumption. Conversely, the consumption of the five major seasonings showed an elastic behaviour. This suggests that, as the food budget increases, the consumption of seasonings increases at a higher rate than the budget itself.

The analysis of uncompensated own-price elasticity reveals contrasting patterns between food and the five major seasoning types. Food consumption is elastic, with price increases resulting in substantial decreases in consumption. However, the data are not sufficiently specific to explain which food categories experienced a decline in consumption. By contrast, the five seasonings show price inelasticity, although they are still negative. This indicates that, while price increases in these commodities lead to reduced consumption, the magnitude of the reduction is proportionally smaller than the price increase. These findings highlight differential consumer responses to price changes across food and seasoning categories.

TABLE 4

ESTIMATION RESULTS FOR INCOME AND PRICE ELASTICITY

Elasticities at different stages	Income elasticity	Own-price elasticity (Marshalian)	Own-price elasticity (Hicksian)
A. Food expenditure with respect to total income (First stage—unconditional)	0.741***	-1.063***	-0.687***
	(0.020)	(0.081)	(0.077)
B. Five seasonings consumption with respect to food expenditure (Second stage—conditional)	1.117***	-0.495***	-0.426***
	(0.064)	(0.045)	(0.045)
C. Specific seasonings consumption with respect to five consumption seasonings (Third stage—conditional)			
Cooking oil	0.903***	-0.950***	6.607
	(0.028)	(0.072)	(34.456)
Red onion	0.892***	-0.939***	-3.705
	(0.038)	(0.291)	(19.916)
Garlic	0.899***	-0.999***	7.34
	(0.014)	(0.128)	(43.604)
Red chili	0.897***	-0.969***	-22.031
	(0.004)	(0.304)	(103.212)
Cayenne pepper	0.900***	-0.986***	7.945
	(0.015)	(0.077)	(44.737)

Source: Data processed.

Notes: *, **, and *** refer to the significant level in 10%, 5%, and 1%, respectively.
(…) refers to standard deviation.

The parameters for each commodity in Table 4 are conditional elasticities. Table 5 provides a comparative analysis of these results with the unconditional elasticities. In the previous table, all the compensated (Hicksian) elasticities were statistically insignificant, making it difficult to calculate the unconditional (Marshallian) elasticity. To address this issue, I use equation (11) to obtain significant values for compensated elasticity. Columns 2-4 of Table 5 present the conditional elasticities obtained from stage 3, while columns 5-7 present the unconditional elasticities that combine the estimation results from stages 1, 2, and 3.

TABLE 5

CONDITIONAL VS UNCONDITIONAL ELASTICITY

Commodities	**Conditional**			**Unconditional**		
	Income elasticity	Own-price (Marshalian)	Own-price (Hicksian)	Income elasticity	Own-price (Marshalian)	Own-price (Hicksian)
Cooking oil	0.903	-0.950	-0.599	0.747	-0.774	-0.765
Red onion	0.892	-0.939	-0.768	0.738	-0.853	-0.849
Garlic	0.899	-0.999	-0.886	0.744	-0.942	-0.940
Red chili	0.897	-0.969	-0.850	0.742	-0.910	-0.906
Cayenne pepper	0.900	-0.986	-0.841	0.745	-0.913	-0.910

Source: Data processed.
Notes: Own-price elasticity (Hicksian/compensated) are calculated using equation (11).

Unconditional elasticity is the primary coefficient employed in the interpretation and simulation of the consumption responses. A comparison between conditional and unconditional elasticities reveals that unconditional income elasticities are 0.1–0.2 points lower than conditional income elasticities, reflecting the parameter correction in the previous stage. Moreover, the analysis indicates that both Marshallian and Hicksian unconditional own-price elasticities yield similar results, suggesting that the substitution effect due to price changes is substantially higher than the income effect—Marshallian elasticity includes income and substitution effects, while Hicksian elasticity excludes the income effect. Furthermore, the same table shows that both income and price elasticity are inelastic. Income elasticities across the five commodities show similar patterns. However, the lowest level of own-price elasticity was observed for cooking oil among the commodities studied.

Additionally, this study examines the potential substitution-complement relationships among seasonings. In the third stage, both uncompensated and compensated cross-price elasticities were calculated. As Table 6 shows, the results indicate that none of the coefficients are statistically significant. Consequently, the presence of substitution or complementarity among the five seasoning commodities can be ignored.

TABLE 6

CONDITIONAL CROSS PRICE ELASTICITY IN STAGE THREE

Commodities	**Cooking oil**	**Red onion**	**Garlic**	**Red chili**	**Cayenne pepper**
			Marshalian/Uncompensated		
Cooking oil	-	0.024	-0.007	0.033	-0.003
	-	(0.054)	(0.163)	(0.300)	(0.150)
Red onion	0.028	-	-0.029	0.083	-0.035
	(0.190)	-	(0.325)	(0.627)	(0.374)
Garlic	0.021	0.034	-	0.045	0.000
	(0.197)	(0.095)	-	(0.357)	(0.137)
Red chili	0.032	0.029	0.000	-	0.010
	(0.146)	(0.075)	(0.135)	-	(0.099)
Cayenne pepper	0.024	0.036	-0.002	0.030	-
	(0.184)	(0.103)	(0.144)	(0.288)	-
			Hicksian/Compensated		
Cooking oil	-	-2.776	8.364	-21.153	8.959
	-	(19.710)	(43.675)	(103.263)	(44.849)
Red onion	7.493	-	8.240	-20.845	8.817
	(34.689)	-	(43.880)	(103.780)	(45.129)
Garlic	7.551	-2.756	-	-21.063	8.928
	(34.587)	(19.724)	-	(103.260)	(44.802)
Red chili	7.545	-2.755	8.322	-	8.918
	(34.540)	(19.690)	(43.608)	-	(44.763)
Cayenne pepper	7.555	-2.755	8.340	-21.085	-
	(34.577)	(19.737)	(43.625)	(103.193)	-

Source: Data processed.

Notes: *, **, and *** refer to the significant level in 10%, 5%, and 1%, respectively.
(…) refers to standard deviation.

6.3 Heterogeneity Analysis

I conducted a heterogeneity analysis by categorizing the sample based on three criteria: cuisine characteristics, COVID-19 pandemic period, and regional economic size. The first analysis focused on the common culinary profile in each province, with particular emphasis on the tendency towards spicy cuisine. This culinary profile was based on the findings of Wijaya (2020). From 33 provinces studied, 21 were classified as predominantly spicy foods, primarily in Sumatra, Sulawesi, Bali, and some parts of Java. The remaining provinces, including Kalimantan, Maluku, Papua, and other parts of Java, were classified as non-spicy cuisine regions.

Referring to Table 7, the results reveal that provinces characterized by spicy cuisines have higher income elasticity in the first and second stages. In the third stage—conditional elasticity—only the cooking oil and red onions showed higher elasticity in these regions. However, based on the unconditional elasticity—defined in equation (12)—income elasticity tends to be elevated in provinces associated with spicy cuisine. In terms of own-price elasticity, red onions, garlic, and red chili were found to increase significantly in provinces with spicy food preferences. Conversely, cooking oil and cayenne pepper showed lower elasticity than the non-spicy food provinces.

The second heterogeneity analysis compared the periods before and after the COVID-19 pandemic, with the pre-pandemic period defined as 2015h1–2019h2 and the post-pandemic period as 2020h1–2022h2. The results show that there is no significant variation between these two periods, with income and own-price elasticities remaining constant. This consistency suggests that consumer behavior towards these seasonings has remained largely unchanged, despite the pandemic. Moreover, these findings highlight the persistent importance of these five commodities in the Indonesian daily consumption landscape.

TABLE 7
HETEROGENEITY ANALYSIS

Category	Elasticity						
	First stage (Unconditional)	Second stage (Conditional)	Third stage (Conditional)				
			Cooking oil	Red onion	Garlic	Red chili	Cayenne pepper
Cuisines characteristic							
Spicy							
Income elasticity	0.788***	1.200***	0.923***	0.946***	0.908***	0.893***	0.916***
	(0.027)	(0.081)	(0.010)	(0.036)	(0.005)	(0.001)	(0.010)
Own-price (Marshalian)	-1.276***	-0.464***	-0.717***	-1.096***	-1.101***	-1.047***	-0.881***
	(0.106)	(0.054)	(0.169)	(0.179)	(0.128)	(0.039)	(0.078)
Own-price (Hicksian)	-0.875***	-0.388***	0.689	-0.710*	-0.052	-3.788***	0.018
	(0.102)	(0.054)	(0.457)	(0.394)	(0.511)	(0.499)	(0.529)
Other							
Income elasticity	0.641***	1.150***	0.918***	0.913***	0.915***	0.897***	0.939***
	(0.033)	(0.094)	(0.011)	(0.008)	(0.023)	(0.001)	(0.103)
Own-price (Marshalian)	-0.529***	-0.485***	-0.986***	-0.908***	-0.905***	-0.952***	-0.993***
	(0.141)	(0.069)	(0.041)	(0.046)	(0.202)	(0.065)	(0.252)
Own-price (Hicksian)	-0.202	-0.418***	1.063	0.449	-0.006	-4.621**	-0.630
	(0.131)	(0.070)	(1.118)	(0.753)	(1.149)	(1.919)	(0.787)
COVID-19 pandemic							
Before pandemic							
Income elasticity	0.753***	1.105***	0.911***	0.917***	0.914***	0.895***	0.920***
	(0.025)	(0.076)	(0.009)	(0.020)	(0.050)	(0.004)	(0.074)
Own-price (Marshalian)	-1.030***	-0.486***	-0.907***	-1.032***	-0.696	-1.018***	-0.614
	(0.109)	(0.057)	(0.055)	(0.127)	(1.350)	(0.089)	(1.559)
Own-price (Hicksian)	-0.647***	-0.423***	1.990	-0.063	-0.087	-5.173	0.066
	(0.104)	(0.057)	(1.924)	(1.135)	(0.717)	(3.539)	(0.843)

Category	Elasticity						
	First stage (Unconditional)	Second stage (Conditional)	Third stage (Conditional)				
			Cooking oil	Red onion	Garlic	Red chili	Cayenne pepper
During and after pandemic							
Income elasticity	0.684***	1.287***	0.923***	0.914***	0.902***	0.897***	0.904***
	(0.037)	(0.120)	(0.009)	(0.010)	(0.009)	(0.006)	(0.007)
Own-price (Marshalian)	-1.079***	-0.580***	-0.872***	-0.828***	-0.826**	-0.997***	-0.939***
	(0.127)	(0.074)	(0.064)	(0.264)	(0.383)	(0.038)	(0.072)
Own-price (Hicksian)	-0.748***	-0.500***	0.832	0.884	0.471	-6.230***	0.582
	(0.118)	(0.075)	(0.818)	(1.641)	(0.847)	(2.282)	(1.044)
Economy size							
Small economy							
Income elasticity	0.772***	1.252***	0.907***	1.345	1.064	0.892***	0.868***
	(0.027)	(0.071)	(0.013)	(18.939)	(6.106)	(0.015)	(0.252)
Own-price (Marshalian)	-1.158***	-0.497***	-0.872***	-1.893	-0.137	-1.116**	-1.731
	(0.096)	(0.050)	(0.070)	(39.089)	(34.375)	(0.465)	(4.908)
Own-price (Hicksian)	-0.765***	-0.418***	2.939	-1.834	-0.059	-3.592	-2.203***
	(0.093)	(0.050)	(3.738)	(40.744)	(31.951)	(4.407)	(0.720)
Large economy							
Income elasticity	0.634***	0.504***	0.919***	1.053	0.910***	0.899***	0.914***
	(0.019)	(0.100)	(0.010)	(1.150)	(0.011)	(0.003)	(0.018)
Own-price (Marshalian)	-0.553***	-0.597***	-0.921***	-0.192	-0.869***	-1.197***	-0.890***
	(0.104)	(0.066)	(0.055)	(6.106)	(0.189)	(0.188)	(0.066)
Own-price (Hicksian)	-0.226**	-0.564***	0.892	-0.052	0.775	-4.944	0.258
	(0.096)	(0.069)	(0.929)	(5.215)	(1.796)	(3.715)	(1.480)

Source: Data processed.

Notes: *, **, and *** refer to the significant level in 10%, 5%, and 1%, respectively.
(…) refers to standard deviation.

The third heterogeneity analysis categorizes provinces according to their economic size, which is determined by the share of regional GDP in the national GDP. Among the 33 provinces evaluated, eight were identified as having substantial economic shares: Sumatra Utara, Riau, Jakarta, Jawa Barat, Jawa Tengah, Jawa Timur, Banten, and Kalimantan Timur. These provinces were among the top 25 percent of the provinces with the highest GDP shares and met the minimum annual share threshold of 3 percent. The remaining provinces were classified as having small-to-medium-sized economies. The analysis revealed that although the third stage shows no significant difference in income and own-price elasticities, the first and second stages indicate that smaller provinces tend to have higher elasticities. This finding suggests that provinces with smaller economies are more responsive to income and price fluctuations than are those with larger economies.

6.4 Robustness Check

To verify the stability of the elasticity coefficients, a robustness check is conducted in the third stage. I use alternative data sources for the price variable while maintaining the existing model specifications. The benchmark model used price data from a consumer price survey, while the robustness check used price data from Socioeconomic Survey (SUSENAS). Price data were obtained by dividing the consumption values by quantity. Table 8 presents a comparison of the coefficients obtained from both data sources. The results demonstrate that income elasticity remains relatively consistent across data sources, with similar values and levels of significance. However, the uncompensated own-price elasticity shows different results for cooking oil and cayenne peppers. Specifically, the coefficients obtained in the robustness check for these two seasonings were not statistically significant. Notably, despite the lack of statistical significance, the results of the robustness check for cooking oil retained the same negative sign as the benchmark model.

TABLE 8
ROBUSTNESS CHECKS ON STAGE THREE
(CONDITIONAL ELASTICITY)

Commodities	Consumer Price Survey			Socioeconomic Survey (SUSENAS)		
	Income elasticity	Own-price (Marshalian)	Own-price (Hicksian)	Income elasticity	Own-price (Marshalian)	Own-price (Hicksian)
Cooking oil	0.903***	-0.950***	6.607	0.859***	-1.437	-2.285
	(0.028)	(0.072)	(34.456)	(0.318)	(2.993)	(4.160)
Red onion	0.892***	-0.939***	-3.705	0.910***	-0.921***	0.777
	(0.038)	(0.291)	(19.916)	(0.018)	(0.074)	(2.993)
Garlic	0.899***	-0.999***	7.340	0.902***	-0.928***	3.680
	(0.014)	(0.128)	(43.604)	(0.004)	(0.151)	(7.220)
Red chili	0.897***	-0.969***	-22.031	0.897***	-1.026***	-5.705
	(0.004)	(0.304)	(103.212)	(0.004)	(0.137)	(6.313)
Cayenne pepper	0.900***	-0.986***	7.945	0.970	0.615	0.837
	(0.015)	(0.077)	(44.737)	(2.058)	(49.063)	(43.105)

Source: Data processed.
Notes: *, **, and *** refer to the significant level in 10%, 5%, and 1%, respectively.
(…) refers to standard deviation.

6.4 Simulation Analysis

Using the coefficients from previous results, simulations were performed based on three different scenarios. The first scenario simulates a simultaneous 10 percent increase in both income and commodity prices. The second scenario simulates a 10 percent increase in income while holding the prices constant. The third scenario examines the impact of a 10 percent increase in prices with no change in income. For all three scenarios, I use the average per capita consumption data across all provinces for 2022 as the baseline. These simulations allow me to assess the potential impact of changes in income and prices on final consumption.

TABLE 9
CONSUMPTION RESPONSE SIMULATION

Commodities	At original price (consumption in 2022)	Scenario 1	Scenario 2	Scenario 3
Average expenditure per capita in rupiahs				
Cooking oil	19,464.700	19,412.411	20,919.512	17,957.599
Red onion	9,039.333	8,935.403	9,706.712	8,268.024
Garlic	5,607.288	5,496.046	6,024.526	5,078.807
Red chili	7,460.015	7,335.372	8,013.879	6,781.508
Cayenne pepper	8,278.803	8,139.407	8,895.514	7,522.697
Change in average expenditure (percentage)				
Cooking oil	-	-0.269	7.474	-7.743
Red onion	-	-1.150	7.383	-8.533
Garlic	-	-1.984	7.441	-9.425
Red chili	-	-1.671	7.424	-9.095
Cayenne pepper	-	-1.684	7.449	-9.133

Source: Data processed.
Notes: Scenario 1: income increase 10% and price increase 10%.
Scenario 2: income increase 10% and price constant.
Scenario 3: income constant and price increase 10%.
The price elasticity used is Marshalian (uncompensated).
We employed equation (11) to obtain the alternative parameter.

As illustrated in Table 9, the simulation displays changes in consumption in both real monetary units and percentages. The table results demonstrate that price elasticity exceeds income elasticity for all commodities. This relationship is particularly evident in Scenario 1, where a simultaneous increase in income and price results in a net decrease in total consumption. This outcome is attributed to the fact that the reduction in consumption due to price increases significantly outweighs the increase in consumption due to income increases. For instance, in Scenario 1, the red onion consumption decreased by 1.15 percent. Among the five commodities examined, cooking oil demonstrated the smallest percentage decrease at 0.3 percent, which was nearly negligible. In contrast, the remaining four commodities demonstrated more substantial declines in consumption, ranging from 1.15 to 1.98 percent.

7. Discussion And Policy Implication

Research on food consumption has extensively studied various commodities. However, seasoning consumption patterns remain largely unexplored. This study addresses this gap by conducting the first comprehensive investigation of seasoning consumption in Indonesia. Using the Linear Approximate Almost Ideal Demand System (LA/AIDS) model, this study provides significant findings on the elasticity of seasoning consumption with respect to income and price.

Five key seasonings play a central role in Indonesian cuisine: cooking oil, red onions, garlic, red chili, and cayenne pepper. These seasonings are essential components of daily culinary practices and significantly influence the cuisine flavour. Their ubiquity makes them irreplaceable items in the household grocery list. Moreover, the results of the LA/AIDS model indicate that the consumption of these seasonings is a necessity, as evidenced by their consumption elasticity with respect to income. As income rises, the consumption of these seasonings also increases, but at a lower rate than income growth itself. Thus, these results confirm that consumers respond predictably to the market signals. Therefore, although culturally important for daily consumption in Indonesia, demand showed an inelastic but not negligible response.

Among the five main seasonings in Indonesia, this study found that cooking oil is the most important for cuisine flavour, accounting for 39 percent of the total consumption. Red onions emerged as the second most important ingredient, accounting for 19 percent of the total consumption. These two commodities are essential ingredients in various cuisines that significantly enhance the final flavour. It is therefore not surprising that red onions in particular are consumed at a rate of 2.3 ounces per month per capita, almost double than garlic.

Chillies—particularly red chillies and cayenne peppers—are considered strategic commodities in Indonesia owing to their significant consumption share. The consumption share of each commodity exceeds 10 percent. Chili is an important ingredient in all regions, although the amount used may vary depending on local preferences. The importance of chili in the Indonesian diet is further emphasized by the high per capita consumption, which is 0.16 kg per month for each red chili and cayenne pepper. This high consumption rate makes it difficult to eliminate them from cuisine compositions, thus highlighting their role in Indonesian culinary culture (Wijaya et al., 2020).

While there are some similarities among the five commodities—between red onion and garlic, and between red chili and cayenne pepper—they do not reveal any substitution or complementarity relationships with each other. Each commodity plays a unique role in shaping the flavour landscape of Indonesian cuisine, as evidenced by the lack of significant cross-price elasticity coefficients. However, consumers respond to price increases in these seasonings by reducing consumption. For every 10 percent price increase, consumption falls by about 7.7 to 9.4 percent. Notably, this reduction in consumption persists even when income rises, suggesting that the negative effect of price increases on consumption outweighs the positive effect of income growth.

Further analysis of heterogeneity yielded another important result, especially when categorized by cuisine type. Regions characterized by spicy cuisine have higher elasticity coefficients, especially for the income elasticity and own-price elasticity of red onion, garlic, and red chili. Conversely, classifications based on economic size and COVID-19 pandemic periods did not cause significant variations in the elasticity coefficients. These results highlight the importance of regional culinary preferences in shaping seasoning consumption patterns.

As a policy implication, the government possesses various mechanisms to intervene in the market through its own authority. The main goal of these policies is to maintain purchasing power and the level of welfare. Currently, the government regulates the maximum retail prices of certain goods to prevent excessive price increases. Although this approach may seem effective, it can lead to market inefficiencies. Instead of imposing maximum retail prices, the government can stabilize prices by eliminating supply side frictions.

The government has three mechanisms to eliminate this friction. First, regional connectivity can be improved by providing adequate transportation facilities. So far, most of these seasonings have been imported locally from production centres in other provinces. For instance, statistical publications by BPS-Statistics Indonesia show that the main centres of frying oil production are concentrated in Sumatra, Java, and Kalimantan (BPS, 2024). Red onions are mainly produced in Jawa Tengah and Jawa Timur, and red chilies are mainly produced in Jawa Barat and Sumatra Utara (BPS, 2022a, 2022b). Therefore, connectivity and transportation remain crucial, as supported by Shively and Thapa (2017), who found that connectivity accounts for more than 50 percent of the volatility in agricultural commodity prices. Reducing transportation costs—through improved connectivity—can lead to lower food prices (Liu et al., 2025).

Second, the government must mitigate disruptions in the food supply chain caused by production delays among local farmers (Hommes et al., 2022). Anticipating crop failures due to pests, diseases, and natural disasters remains critical as they directly lead to shortages and price hikes in the market. Additionally, it is necessary to optimize the harvesting period by spreading out the harvest time and avoiding concentration during certain periods. This will prevent oversupply at certain times and minimize the risk of shortages at others. Third, when domestic agricultural production fails to meet market demand, the government should adopt flexible import policies (Shobur et al., 2025), including reducing import barriers through quotas and tariffs. This approach will help keep the supply of seasonings in the domestic market stable and minimize potential price fluctuations.

8. Concluding Remark

The empirical results confirm the existing demand theory and show inelastic demand elasticity for seasonings, with respect to both income and price. As expected, consumption rises with increasing income and falls as prices rise. Although seasonings seem to be essential goods for public consumption, consumer behaviour aligns with the market mechanisms. In particular, the results indicate that price increases have a more pronounced negative effect on reducing consumption quantity than the positive effect of rising income. Given the relatively inelastic demand for essential seasonings, policy responses should focus on promoting competitive supply conditions and minimizing market frictions rather than imposing direct price controls. This approach would preserve consumer welfare and minimize potential market distortions.

Furthermore, this study offers broader implications for managing food security in other countries. While this study limited the focus to certain seasonings, these commodities are still known to be essential for daily consumption. I realize that each country has its own consumption pattern, including which commodities are important to their people. Referring to my study, I found that although these commodities are still essential and frequently consumed, their consumption response still follows the market signals. Based on this, I suggest to other countries that are highly dependent on imports—such as those in Africa, the Middle East, and East Asia that rely on imports for their domestic food consumption—should identify which commodities remain important but vulnerable. These countries are often the most vulnerable to international food trade shocks. Therefore, it is necessary to mitigate resilience and food security earlier by creating a multiple food supply chain, possibly through domestic production, alternative trading partners, and substitution commodities.

Although my study successfully obtained unconditional elasticity estimates along with heterogeneity analysis, this study has certain limitations that should be addressed in the future. First and foremost, my reliance on macro-level data and the lack of access to micro-level datasets limit the ability to conduct an in-depth analysis. I recommend that future studies should examine the micro data level to gain a more nuanced understanding of consumption patterns across income levels—lower, middle, and upper classes. In addition, by using micro-level datasets, researchers can further explore the extent to which people will continue to increase their consumption of seasonings to enhance their food flavour. This approach will provide policymakers with sharper insights into the impact of price increases on overall consumption.

The second limitation relates to the use of static demand models. Although this study used instrumental variables to address endogeneity issues related to price and consumption variables, still it did not include dynamic LA/AIDS models. The use of a dynamic model would allow researchers to decompose the elasticity into short-run and long-run effects. However, obtaining these coefficients in STATA is challenging due to the current unavailability of appropriate commands. Future research could make a significant contribution by developing a STATA command—potentially published in The STATA Journal—that simultaneously addresses three key aspects: implementing instrumental variables, accounting for new product entry, and estimating both short- and long-run parameters. Such an approach would provide unbiased estimates and would better reflect the real-world conditions.

The third limitation concerns the threshold for seasoning consumption. While seasonings are essential in cuisines and have become a basic necessity, there are likely lower and upper limits on per capita consumption. These limits are influenced by individual dietary patterns and the ideal measurements of cuisine recipes. This suggests that the excessive consumption of seasonings is not possible. Conversely, even in the context of high seasoning prices, it is unlikely that consumption will be completely eliminated or significantly reduced—instead, it is likely to remain at the baseline level. Again, it was not possible in my study to determine the upper and lower thresholds for the consumption level. Therefore, future studies using experimental methods are needed to accurately determine these thresholds.

References

Abidin, M. R. Z., Che Ishak, F. A., Ismail, I. A., & Juhari, N. H. (2020). Modern Malaysian cuisine: Identity, culture, or modern-day fad? *International Journal of Gastronomy and Food Science*, 21, 100220. https://doi.org/10.1016/j.ijgfs.2020.100220

Bakhtavoryan, R., & Capps, O. (2024). A demand systems analysis for cheese varieties using a balanced panel of US-designated market areas, 2018–2020. *Journal of Agricultural and Resource Economics*, 49(2), 203–220. https://doi.org/10.22004/AG.ECON.338990

Basak, S., Chakraborty, S., & Singhal, R. S. (2023). Revisiting Indian traditional foods: A critical review of the engineering properties and process operations. *Food Control*, *143*, 109286. https://doi.org/10.1016/j.foodcont.2022.109286

Bilgic, A., & Yen, S. T. (2013). Household food demand in Turkey: A two-step demand system approach. *Food Policy*, 43, 267–277. https://doi.org/10.1016/j.foodpol.2013.09.004

Blanciforti, L., & Green, R. (1983). An almost ideal demand system incorporating habits: An analysis of expenditures on food and aggregate commodity groups. *The Review of Economics and Statistics*, 65(3), 511. https://doi.org/10.2307/1924200

Bonke, J., & Browning, M. (2009). The allocation of expenditures within the household: A new survey. *Fiscal Studies*, 30(3–4), 461–481. https://doi.org/10.1111/j.1475-5890.2009.00104.x

Bouyssou, C. G., Jensen, J. D., & Yu, W. (2024). Food for thought: A meta-analysis of animal food demand elasticities across world regions. *Food Policy*, 122, 102581. https://doi.org/10.1016/j.foodpol.2023.102581

BPS. (2022a). *Distribution flow of red chili in Indonesia 2022* (Statistic Publication 06100.2253). BPS-Statistics Indonesia. https://www.bps.go.id/id/publication/2022/10/24/1c3eb22e94bf4523bc029d1b/distribusi-perdagangan-komoditas-cabai-merah-di-indonesia-2022.html

BPS. (2022b). *Distribution flow of red onion in Indonesia 2022* (06100.2254). BPS-Statistics Indonesia. https://www.bps.go.id/id/publication/2022/10/24/958ef61ffcb0e88357bb99d1/distribusi-perdagangan-komoditas-bawang-merah-di-indonesia-2022.html

BPS. (2022c). *Official statistics news*. BPS-Statistics Indonesia.

BPS. (2024). *Distribution flow of frying oil in Indonesia 2023* (Statistic Publication 06100.2359). BPS-Statistics Indonesia. https://www.bps.go.id/id/publication/2023/11/22/c7a34094ff453f2a0f8dc7b8/distribusi-perdagangan-komoditas-minyak-goreng-indonesia-2023.html

Burke, P. J., & Abayasekara, A. (2018). The price elasticity of electricity demand in the United States: A three-dimensional analysis. *The Energy Journal, 39*(2), 123–146. https://doi.org/10.5547/01956574.39.2.pbur

Casabianca, M. S., Gallego, J. M., Góngora, P., & Rodríguez-Lesmes, P. (2022). Price elasticity of demand for voluntary health insurance plans in Colombia. *BMC Health Services Research, 22*(1), 618. https://doi.org/10.1186/s12913-022-07899-2

Chai, A., Kiedaisch, C., & Rohde, N. (2024). Household spending diversity, aggregation, and the value of product variety. *Economic Modelling*, *141*, 106857. https://doi.org/10.1016/j.econmod.2024.106857

Čiderová, T., & Ščasný, M. (2022). Estimation of alcohol demand elasticity: Consumption of wine, beer, and spirits at home and away from home. *Journal of Wine Economics*, *17*(4), 329–337. https://doi.org/10.1017/jwe.2022.42

Colen, L., Melo, P. C., Abdul-Salam, Y., Roberts, D., Mary, S., & Gomez Y Paloma, S. (2018). Income elasticities for food, calories and nutrients across Africa: A meta-analysis. *Food Policy*, *77*, 116–132. https://doi.org/10.1016/j.foodpol.2018.04.002

Deaton, A., & Muellbauer, J. (1980). An Almost Ideal Demand System. *The American Economic Review*, *70*(3), 312–326.

Deryugina, T., MacKay, A., & Reif, J. (2020). The long-run dynamics of electricity demand: Evidence from municipal aggregation. *American Economic Journal: Applied Economics*, *12*(1), 86–114. https://doi.org/10.1257/app.20180256

Dey, M. M., Alam, Md. F., & Paraguas, F. J. (2011). A multi-stage budgeting approach to the analysis of demand for fish: An application to inland areas of Bangladesh. *Marine Resource Economics*, *26*(1), 35–58. https://doi.org/10.5950/0738-1360-26.1.35

Díaz, A. O., & Medlock, K. B. (2021). Price elasticity of demand for fuels by income level in Mexican households. *Energy Policy*, *151*, 112132. https://doi.org/10.1016/j.enpol.2021.112132

Edgerton, D. L. (1997). Weak separability and the estimation of elasticities in multi-stage demand systems. *American Journal of Agricultural Economics*, *79*(1), 62–79. https://doi.org/10.2307/1243943

Ellis, R. P., Martins, B., & Zhu, W. (2017). Health care demand elasticities by type of service. *Journal of Health Economics*, *55*, 232–243. https://doi.org/10.1016/j.jhealeco.2017.07.007

Farag, M., NandaKumar, A. K., Wallack, S., Hodgkin, D., Gaumer, G., & Erbil, C. (2012). The income elasticity of health care spending in developing and developed countries. *International Journal of Health Care Finance and Economics*, *12*(2), 145–162. https://doi.org/10.1007/s10754-012-9108-z

Fleissig, A. R. (2021). Expenditure and price elasticities for tourism sub-industries from the fourier flexible form. *Tourism Economics*, *27*(8), 1692–1706. https://doi.org/10.1177/1354816620938170

Fouquet, R. (2014). Long-run demand for energy services: Income and price elasticities over two hundred years. *Review of Environmental Economics and Policy*, *8*(2), 186–207. https://doi.org/10.1093/reep/reu002

Fredman, P., & Wikström, D. (2018). Income elasticity of demand for tourism at Fulufjället national park. *Tourism Economics*, *24*(1), 51–63. https://doi.org/10.1177/1354816617724012

Gatt, W., & Falzon, J. (2014). British tourism demand elasticities in Mediterranean countries. *Applied Economics*, *46*(29), 3548–3561. https://doi.org/10.1080/00036846.2014.934432

Goetzke, F., & Vance, C. (2021). An increasing gasoline price elasticity in the United States? *Energy Economics*, *95*, 104982. https://doi.org/10.1016/j.eneco.2020.104982

Hamzah, I. N., & Huang, W. (2023). The dynamics of strategically important food preference in Indonesia: An empirical evaluation of consumption pattern and welfare loss. *Economic Analysis and Policy*, *79*, 435–449. https://doi.org/10.1016/j.eap.2023.06.024

Hasiner, E., & Yu, X. (2020). Meat Consumption and Democratic Governance: A Cross-National Analysis. *China Economic Review*, *59*, 1–13. https://doi.org/10.1016/j.chieco.2016.06.008

Helgy Library. (2023, October). *Spice consumption (total)*. Spice Consumption (Total). https://www.helgilibrary.com/indicators/spice-consumption-total

Hommes, C., Li, K., & Wagener, F. (2022). Production delays and price dynamics. *Journal of Economic Behavior & Organization*, *194*, 341–362. https://doi.org/10.1016/j.jebo.2021.12.033

Jin, T., & Kim, J. (2022). The elasticity of residential electricity demand and the rebound effect in 18 European union countries. *Energy Sources, Part B: Economics, Planning, and Policy*, *17*(1), 2053896. https://doi.org/10.1080/15567249.2022.2053896

Keane, M., & Neal, T. (2023). Instrument strength in IV estimation and inference: A guide to theory and practice. *Journal of Econometrics*, *235*(2), 1625–1653. https://doi.org/10.1016/j.jeconom.2022.12.009

Labandeira, X., Labeaga, J. M., & López-Otero, X. (2017). A meta-analysis on the price elasticity of energy demand. *Energy Policy*, *102*, 549–568. https://doi.org/10.1016/j.enpol.2017.01.002

Lecocq, S., & Robin, J.-M. (2015). Estimating Almost-ideal Demand Systems with Endogenous Regressors. *The Stata Journal*, *15*(2), 554–573. https://doi.org/10.1177/1536867X1501500214

Liu, X., Zhang, Y., Lan, X., & Si, W. (2025). Can reducing agricultural trade costs foster the transformation of the agrifood system? Evidence from China. *China Economic Review*, *91*, 102406. https://doi.org/10.1016/j.chieco.2025.102406

Marioni, L. da S., Aznar, A. R., Aitken, A., Kapur, S., Smith, R., & Beckert, W. (2022). *Estimating food and drink elasticities*. National Institute of Economic and Social Research. https://www.niesr.ac.uk/wp-content/uploads/2021/11/Estimating-Food-and-Drink-Demand-Elasticities.pdf

Nguyen, L., Duong, L. T., & Mentreddy, R. S. (2019). The U.S.import demand for spices and herbs by differentiated sources. *Journal of Applied Research on Medicinal and Aromatic Plants*, *12*, 13–20. https://doi.org/10.1016/j.jarmap.2018.12.001

Pakes, A. (2021). A helicopter tour of some underlying issues in empirical industrial organization. *Annual Review of Economics*, *13*(Volume 13, 2021), 397–421. https://doi.org/10.1146/annurev-economics-082019-023513

Parappurathu, S., & Mathur, C. V. (2006). Analysis of demand for major spices in India. *Agricultural Economics Research Review*, *19*, 367–376.

Pollak, R. A., & Wales, T. J. (1969). Estimation of the linear expenditure system. *Econometrica*, 37(4), 611. https://doi.org/10.2307/1910438

Ramírez, A. (2013). A multi-stage almost ideal demand system: The case of beef demand in Colombia. *Revista Colombiana de Estadística*, *36*(1), 23–42.

Rathnayaka, S. D., Selvanathan, S., & Selvanathan, E. A. (2021). Demand for Animal-Derived Food in Selected Asian Countries: A System-Wide Analysis. *Agricultural Economics*, *52*(1), 97–122. https://doi.org/10.1111/agec.12609

Rathnayaka, S. D., Selvanathan, S., Selvanathan, E. A., & Kler, P. (2019). Modelling Sri Lankan consumption patterns using error corrected LA-AIDS. *Economic Modelling*, *80*, 185–191. https://doi.org/10.1016/j.econmod.2018.11.006

Roosen, J., Staudigel, M., & Rahbauer, S. (2022). Demand elasticities for fresh meat and welfare effects of meat taxes in Germany. *Food Policy*, *106*, 102194. https://doi.org/10.1016/j.foodpol.2021.102194

Sanderson, E., Spiller, W., & Bowden, J. (2021). Testing and correcting for weak and pleiotropic instruments in two-sample multivariable Mendelian randomization. *Statistics in Medicine*, *40*(25), 5434–5452. https://doi.org/10.1002/sim.9133

Selvanathan, S., Jayasinghe, M., Selvanathan, E. A., & Rathnayaka, S. D. (2024). Dynamic modelling of consumption patterns using LA-AIDS: A comparative study of developed versus developing countries. *Empirical*

Economics, *66*(1), 75–135. https://doi.org/10.1007/s00181-023-02465-z

Seya, H., Asaoka, T., Chikaraishi, M., & Axhausen, K. W. (2024). Estimatingthe price elasticity of demand for off-street parking in Hiroshima city, Japan. *Transportation Research Part A: Policy and Practice*, *183*, 104051. https://doi.org/10.1016/j.tra.2024.104051

Sherman, P. W., & Billing, J. (1999). Darwinian gastronomy: Why we use spices. *BioScience*, *49*(6), 453–463. https://doi.org/10.2307/1313553

Shively, G., & Thapa, G. (2017). Markets, transportation infrastructure, and food prices in Nepal. *American Journal of Agricultural Economics*, *99*(3), 660–682. https://doi.org/10.1093/ajae/aaw086

Shobur, M., Nyoman Marayasa, I., Bastuti, S., Muslim, A. C., Pratama, G. A., & Alfatiyah, R. (2025). Enhancing food security through import volume optimization and supply chain communication models: A case study of East Java's rice sector. *Journal of Open Innovation: Technology, Market, and Complexity*, *11*(1), 100462. https://doi.org/10.1016/j.joitmc.2024.100462

Sproesser, G., Ruby, M. B., Arbit, N., Akotia, C. S., Alvarenga, M. D. S., Bhangaokar, R., Furumitsu, I., Hu, X., Imada, S., Kaptan, G., Kaufer-Horwitz, M., Menon, U., Fischler, C., Rozin, P., Schupp, H. T., & Renner, B. (2022). Similar or different? Comparing food cultures with regard to traditional and modern eating across ten countries. *Food Research International*, *157*, 111106. https://doi.org/10.1016/j.foodres.2022.111106

Srivastava, S. K., Kumar, R., Hema, M., & Hasan, R. (2013). Inter-regional variations and future household demand and production of major spices in India. *Journal of Spices and Aromatic Crops*, *22*(1), 47–54.

Stock, J. H., & Watson, M. W. (2020). *Introduction to econometrics* (Global ed.). Pearson.

Suárez-Varela, M. (2020). Modelling residential water demand: An approach based on household demand systems. *Journal of Environmental Management*, *261*, 109921. https://doi.org/10.1016/j.jenvman.2019.109921

Surya, R., & Tedjakusuma, F. (2022). Diversity of sambals, traditional Indonesian chili pastes. *Journal of Ethnic Foods*, *9*(1), 25. https://doi.org/10.1186/s42779-022-00142-7

Untong, A., Ramos, V., Kaosa-Ard, M., & Rey-Maquieira, J. (2014). Thailand's long-run tourism demand elasticities. *Tourism Economics*, *20*(3), 595–610. https://doi.org/10.5367/te.2013.0280

Ventura, R. V., Fernandes, E., Aprigliano Fernandes, V., Cabo, M., Cesar Fadel, A., & Caixeta, R. (2022). Consumer behaviour analysis based on income and price elasticities: The case of the air transportation in

Brazil. *Case Studies on Transport Policy*, *10*(2), 1262–1272. https://doi.org/10.1016/j.cstp.2022.04.012

Wang, S., & Noland, R. B. (2021). What is the elasticity of sharing a ridesourcing trip? *Transportation Research Part A: Policy and Practice*, *153*, 284–305. https://doi.org/10.1016/j.tra.2021.09.008

Wang, Y., & Çakır, M. (2020). Welfare impacts of new demand-enhancing agricultural products: The case of Honeycrisp apples. *Agricultural Economics*, *51*(3), 445–457. https://doi.org/10.1111/agec.12564

Wardman, M. (2024). Investigating demand models with more flexible elasticity functions: Empirical insights from rail demand analysis. *Transportation*. https://doi.org/10.1007/s11116-024-10462-z

Widarjono, A., & Rucbha, S. M. (2016). Household food demand in Indonesia: A two-stage budgeting approach. *Journal of Indonesian Economy and Business*, *31*(1), 163. https://doi.org/10.22146/jieb.15287

Wijaya, C. H., Harda, M., & Rana, B. (2020). Diversity and potency of capsicum spp. Grown in Indonesia. In A. Dekebo (Ed.), *Capsicum*. IntechOpen. https://doi.org/10.5772/intechopen.92991

Wongmonta, S. (2022). An assessment of household food consumption patterns in Thailand. *Journal of the Asia Pacific Economy*, *27*(2), 289–309. https://doi.org/10.1080/13547860.2020.1811191

Wu, B., Shang, X., & Chen, Y. (2021). Household dairy demand by income groups in an urban Chinese province: A multi-stage budgeting approach. *Agribusiness*, 37(3), 629–649. https://doi.org/10.1002/agr.21681

Zhou, J., Xin, X., Li, W., Ding, H., Yu, S., & Cui, X. (2024). Flavor analysis and region prediction of Chinese dishes based on food pairing. *Information Processing & Management*, *61*(3), 103684. https://doi.org/10.1016/j.ipm.2024.103684

Zhou, Z., Su, Y., Gao, J., Xu, L., & Zhang, Y. (2011). New estimates of elasticity of demand for healthcare in rural China. *Health Policy*, *103*(2–3), 255–265. https://doi.org/10.1016/j.healthpol.2011.09.005

Zhuang, R., & Abbott, P. (2007). Price elasticities of key agricultural commodities in *China. China Economic Review*, *18*(2), 155–169. https://doi.org/10.1016/j.chieco.2006.02.006

MENSSAGE PRODUCCIONES
www.menssage.cl

Instrucciones para los autores

Los artículos deberán ser escritos en inglés preferentemente o español y enviados en triplicado a los editores de *Estudios de Economía*.

Los artículos seleccionados serán publicados en el idioma original. Los artículos deberán ser mecanografiados de un solo lado y escritos a doble espacio con márgenes amplios.

- La primera página debe incluir la siguiente información: el título del artículo (debe venir con letras minúsculas en inglés y español), nombre y filiación del autor o autores y un resumen de no más de 100 palabras. Incluya la clasificación del JEL y hasta 5 palabras claves que describan el artículo.
- Las notas de pie de página deberán restringirse al mínimo posible en su cantidad y extensión, ser numeradas correlativamente y evitar incluir fórmulas matemáticas.
- Las fórmulas matemáticas deberán ser numeradas en forma consecutiva. Cuando la derivación de una fórmula ha sido abreviada se deberá presentar en una hoja aparte la derivación del resultado, paso a paso.
- Los cuadros y gráficos se presentarán en hoja aparte. Deberán ser numerados, incluir toda la información necesaria para su interpretación así como la fuente de datos.
- Las referencias bibliográficas deben ser citadas en el texto como sigue: Hirshleifer (1958). Cuando 3 o más autores sean citados deberá hacerse de la siguiente manera: Judge *et al*. (1985). El autor es responsable de que exista una perfecta correspondencia entre el nombre y el año de referencia y las anotaciones de la bibliografía.
- La bibliografía, en orden alfabético, debe listarse al final del artículo de acuerdo al siguiente patrón:
  Hirshleifer, J. (1958). "On the Theory of Optimal Investment Decisions", *Journal of Political Economy*, Vol. 56; 279-394.
  Judge, G., W. Griffiths, R. Carter, H. Lutkepohl y T. Lee. (1985). *The Theory and Practice of Econometrics.* 2ª edición. John Wiley.

Instructions to Authors

Papers should be written in English preferably or Spanish, three copies of the original manuscript should be addressed to Editors of *Estudios de Economía*.

The articles approved will be published in the original language. All submitted manuscripts should be typed on one side, with wide margins and double spaced.

- The first page must include the following information: paper title (should come with lowercase letter's), name, address and affiliation of the author(s) and a summary of no more than 100 words. Include the JEL classification and up to five keywords that describe your paper for indexing purposes.
- Footnotes should be the fewest possible in quantity and extension. They must be numbered in serial order. Please avoid formulas in footnotes.
- Equations should be numbered in consecutive order. When the derivation of a formula has been condensed, an additional page should be included with a full derivation.
- Tables and graphs must be included in a different page. They should be numbered and should include all the information needed for interpretation as well as the data sources.
- References should be quoted in the text as follows: Hirshleifer (1958). Where three or more authors are listed in the reference list, please cite in the text as: Judge *et al*. (1985). The author is responsible for the existence of a perfect correspondence between names and year of the reference and what appear in the section of references.
- References should listed at the end of the paper alphabetically in the following style:
  Hirshleifer, J. (1958). "On the Theory of Optimal Investment Decisions", *Journal of Political Economy*, Vol. 56; 279-394.
  Judge, G., W. Griffiths, R. Carter, H. Lutkepohl and T. Lee. (1985). *The Theory and Practice of Econometrics.* 2nd edition. John Wiley.