The Cliometrics of Onomastics: Modeling Who's Who in Ancient Greece

Laurent Gauthier

The Cliometrics of Onomastics: Modeling Who's Who in Ancient Greece*

Laurent Gauthier†

October 07, 2021

Abstract

In ancient Greece, people carried a single name, and some names were quite common, while others were very rare. If these names were used to distinguish people, why didn't everybody have a different name? Looking into the manner in which the ancient Greeks picked names, we develop an economic model for the existence of names, as a way of exchanging identification information. Considering different information frameworks, we justify the exchange of names in this context as the best system in order to promote cooperation. We then study the optimal choice of names in these conditions and show that the impact of strategic naming on the distribution of names works as an alteration of existing mean-field approaches to name dynamics, which converge to power laws. Strategic naming adds a degree of freedom in the relationship between the observed number of names and the shape of the power law distribution. Confronting these results to empirical data from the archaic and classical periods, we observe that a form of conformist strategic naming could account for the particular shape of the name distribution in Ancient Greece, which differs from contemporary data.

Keywords: Ancient Greece, prisonner's dilemma, onomastics, name distributions, power laws

***Draft working paper** - comments welcome! I wish to thank Karine Karila-Cohen, LAHM and CReAAH (UMR 6566) Université de Rennes, for many discussions on ancient Greek onomastics.

†Email: laurent.o.gauthier@gmail.com. Laboratoire d'Économie Dionysien, Université Paris 8 Saint-Denis Vincennes, EA 3391 - Bâtiment D, 2, rue de la Liberté, 93526 Saint-Denis, France.

Some ancient Greek names have acquired worldwide fame, and are nowadays still given to children. Achilles, Hektor, Demosthenes, or Alexandros are firmly inscribed in European history and literature. These famous names were often given to children already in ancient times, and there are hundreds of epigraphic sources referring to some Achileus or Demosthenes across *Hellas*. Many other names, scarcer, were also given, of which we have only found a handful of instances in the epigraphic or literary record. In classical Greece, there were no family names as we know them today, but single names, sometimes associated to a political group (such as a deme in Athens), and as a result, for all matters and purposes, some people had the same name as many others, while others did not. Why was there such variability in the nature and effective function of the ancient Greeks' names? If unique names were helpful in identifying people, why did not everyone carry a different name? To understand how names were given, we will need to explore how they were used.

Onomastics, the study of names, and prosopography, the study of individual life histories are important facets of ancient Greek history. Names and individual identification have been used to more precisely follow the history of members of the elite, as Puech (2012), Karila-Cohen (2017) and Karila-Cohen (2019) have done. Studying names and how they were attributed has also allowed historians to better grasp the extent of families, as in Bresson (1981), Herman (1990) and Bresson (2019). Using names from epigraphic sources may also help in measuring the reliability of some historical accounts, as Hornblower (2010) proposed to do in the case of Thucydides.

As was discussed in Gauthier (2021), there has been a very substantial effort with the LGPN to collect, and later digitalize, data from hundreds of thousands of inscriptions relating to names in the ancient Greek world. Thanks to the availability of this data, some research has focused on the network aspect of the family relationships based on the names in the inscriptions, such as Cline (2020), Karila-Cohen (2016) and Karila-Cohen (2018). Nevertheless, the historical approach has tended to focus on particular areas, or on particular families, rather than consider names at a much larger scale[1]. Historical approaches have not either looked specifically into the reasons why some names may have been more common than others.

Our goal in this paper is to apply an economic approach to the understanding of ancient Greek names, in order to address both the scale of data, as well as the logic for distributional features of this data. We will hence focus on accounting for the distribution of names, as well as naming strategies. More precise uses of names could be studied, in particular by focusing on their

[1]As indicated in Gauthier (2021), obtaining the LGPN data in bulk is quite difficult, and this may explain the scarcity of research that could rely on the entire dataset.

semantics[2], but the first step consists of analyzing these ancient Greek names in bulk, which has not been done to date, either from a theoretical nor empirical perspective.

We may wonder what economics could have to say about something as un-economic sounding as a name. As a matter of fact, the notion of identity in economics spans a wide range of related concepts. Berg (2019) gives an overview, and for our purposes we can determine that there are three connected axes in this body of research with applications or concepts that are useful for onomastics: group behavior (in the sense of club goods), identity stricto sensu, and conformity. Carr and Landa (1983) considered the benefits of belonging to a group in reducing transaction costs through the enforcement of certain rules, and imposing costs on breaching them. They showed with a simple model how one could account for clans, symbols or family and name as club goods. Depending on various conditions on the cost of informing other group members and the cost of contract breaches, they established an equilibrium for the optimal club size. While the way in which the groups are defined is related to notions of identity, they did not specifically look into defining identity from an economic standpoint. A precise definition of identity was given by Akerlof and Kranton (2000), who proposed to extend a basic utility function to account for their notion of identity: people benefit more from partaking in activities that are line with the prescriptions of their chosen identity. They derived various equilibria, where varying shares of the population choose a particular identity. Not following the prescriptions of one's identity is costly to oneself, because, as they argue, rules are internalized psychologically, and breaching them creates anxiety. Breaching identity-prescribed rules is also costly to others: by observing someone's breaking the rules it arouses emotions that were suppressed in order to internalize the rules in the first place. Their model explains behaviors that could otherwise appear anti-economic. The notions of identity and group belonging are also strongly connected with conformity, which was studied by Bernheim (1994) and extended and formally clarified by Gillen (2015). In this approach, the utility functions of agents of various types include a term for status, which at equilibrium is itself a function of type. In certain conditions, when status plays an important role in utility, then there is an endogenous pooling equilibrium where many agents follow the same action, while only those with extreme preferences have incentives to stray (and receive a lower status). Some experiments have been carried out and showed that conformity affected the outcome of public goods games, as reported by Carpenter (2004).

Names have also been extensively studied in statistical physics. Several publications following

[2] Adressing methodological issues in the study of names, Motschenbacher (2020) has called for the study of names in the context of their use, that is from a linguistics perspective, more than simply studying their frequencies and differences.

physical sciences inspired methods have indeed examined the distribution of names, whether they be family names or given names. As with many social phenomena, the patterns of family name distributions match power laws. Zanette and Manrubia (2001) showed that these distributions exhibited an exponent of 1 (at the density level), and proposed a simple random growth model accounting for this behavior. Baek, Kiet, and Kim (2007) and Maruvka, Shnerb, and Kessler (2010) applied population dynamics models in order to account for the growth in number and distribution of family names. The first paper shows that the shape of family names distributions (generally power laws) can be explained by the way the rate of introduction of new names is expressed. Some other studies carried out purely empirical analyzes of some specific aspects of names distributions, such as dynamics or correlation, in particular Hahn and Bentley (2003), Mateos, Longley, and O'Sullivan (2011), Barucca et al. (2015) and Lee et al. (2016), who looked at given names. Following more of a psychological perspective, Busse and Seraydarian (1977), Young et al. (1993), and Rogerson (2016) focused on some of the implied identity in names. For example, the experiments conducted by Young et al. (1993) showed that people of a younger generation tended to associate first names from an older generation to lesser intelligence and popularity.

In order to tackle the analysis of ancient Greek names, we will first examine the historical context in which names were used at the time. We then develop a model for the optimal behavior that people will adopt with each other as a function of the unique information they have about their counterparty: their name. Considering different information frameworks, we justify the exchange of names in this context as the best system in order to promote cooperation. Then, we will be in a position, relying on that model, to study the optimal choice of names for children, given that equilibrium. With some simplifications, we show that the impact of strategic naming on the distribution of names works as an alteration of the typical mean-field approaches to name dynamics, which converges to a power law. Strategic naming adds a degree of freedom in the relationship between the observed number of names and the shape of the power law distribution. Finally, we confront these reulsts with empirical data from the archaic and classical periods. We show that a form of conformist strategic naming could account for the particular shape of the names' power law distribution in Ancient Greece, which has a parameter around 1.3. This shape is in stark constrats with what has been observed in the large majority of empirical studies on contemporary name distributions, where the parameter was either close to 2 for most countries, or at 1 for Korea.

1 Ancient Greek Names in Context

We begin by addressing the issue of the effect of child mortality on the surviving literary and archaeological traces of names, and then discuss the manner in which names were used and exchanged in ancient Greece.

One potential issue in understanding name choices in ancient Greece is indeed whether *necronymy*, giving a child the name of a prior child who had passed, was common or not. The very high child mortality compared to modern standards, depending on the practice of necronymy, could contribute to a randomization of name choices. The epigraphic sources amost exclusively refer to adults or adolescents, and as a result the names we get to observe are not an unbiased sample of those that were given, because of child mortality. There are historical and cliometric methodological issues in determining the shape of life tables applicable to ancient history, as Woods (2007) summarized. While there is no large amount of direct data that can be exploited, certain extrapolations across time and cultures allow for the linking of life expectancy with fairly stable life table shapes. In summary, one can estimate death rates of around 30% in the first year, and around 50% cumulatively for the first 5 years[3]. Parkin (2013) finds numerous references in classical literature to the dangers of early childhood, and also points out that early infant mortality could reach up to 50% in high stress periods[4]. Such high rates of mortality could naturally interact with name choices. There is no ancient testimony about the practice of necronymy, and whether this practice was common or not would substantially affect the distribution of observed names, since the choices of names that we observe on inscriptions about adults would become randomized. However, according to Hardie (1923), necronymy was common in certain conditions in Macedonia in the early 20th century, and she infers this was most likely inherited from ancient times. In Bresson (1981), comparable observations were made, also pointing towards a fairly systematic use of necronymy. As a result, it is reasonable to assume that despite child mortality, the names of adolescents or adults that we are able to observe today reflect the parents' initial choices, and not a randomized selection.

The extensive analysis of contemporary names we have mentioned earlier may not be relevant for ancient Greece. In addition to the use of a single anthroponym without a patronym in the modern sense, the choice of a name was also strongly affected by tradition and perceived prestige. For example, *paponymy*, the use of a grandparent's name for a child, was frequent. Honor and

[3]See Woods (2007), p. 379-380.

[4]See Parkin (2013), p. 47-48.

names are also closely related. By giving one's name in ancient Greece, a person would identify themselves, and providing this information was a meaningful act. The importance of a name can be shown to be deeply anchored in Greek literature; one good example may be Odysseus's introduction to Alcinous in *Odyssey*, book IX:

> "But now I will tell you my name, so that you too may know it, and then when this pityless day is over I may be your host, although my home is far. I am Odysseus, son of Laertes, known among all men for my cunning, and my fame reaches heaven[5]."

The importance of the name transpires from these verses, but they depict an important encounter between two kings. In contrast, when first meeting someone with whom one is not very well acquainted, names may not be fully given. Odysseus's most famous reply to the Cyclops's question, also in *Odyssey*, book IX, illustrates this rather well. Telling one's name is an important act, and preceeds any form of interaction, and one is supposed to do it in good faith (which is not the case with the Cyclops):

> "My name is Nobody; and my mother, my father and all my companions also call me Nobody[6]"

While certain people may be known by others by reputation, it was not necessarily the case that they knew who the person was. The famous example of Aristides the Just makes this clear: As Plutarch reported[7], he was once confronted in the street by an illiterate man, asking him to write "Aristides" on an ostrakon[8] (Aristides obliged). Whether the story is true is not critically important, but it appears clearly that it was conceivable that a man could know of a name, without being able to recognize the person.

We can consider that the name of a person in classical Greece was used as an important piece of information in dealing with them. As the first step in a relationship, the exchange of names should indeed communicate a certain reputation in some cases, or a degree of prestige, and potentially some clues about the person's origin or family. Cuesta et al. (2015) have shown with behavioral experiments on series of repeated games that reputation acquired in practice played a significant role in human cooperation. When no particular history of interactions is

[5] *Od.*, book IX[16–20] νῦν δ᾽ ὄνομα πρῶτον μυθήσομαι, ὄφρα καὶ ὑμεῖς / εἴδετ᾽, ἐγὼ δ᾽ ἂν ἔπειτα φυγὼν ὕπο νηλεὲς ἦμαρ / ὑμῖν ξεῖνος ἔω καὶ ἀπόπροθι δώματα ναίων. / εἴμ᾽ Ὀδυσεὺς Λαερτιάδης, ὃς πᾶσι δόλοισιν / ἀνθρώποισι μέλω, καί μευ κλέος οὐρανὸν ἵκει.

[6] *Od.*, book IX[366–7] Οὖτις ἐμοί γ᾽ ὄνομα: Οὖτιν δέ με κικλήσκουσι / μήτηρ ἠδὲ πατὴρ ἠδ᾽ ἄλλοι πάντες ἑταῖροι.

[7] Plutarch, *Vitae Parallelae Aristides*, VII.

[8] A pottery fragment used in an ostracism vote, in order to ban a person from Athens.

available, the name may serve as a proxy. The prestige and the reputation are then attached to the names, not to the persons themselves, and reflect the behavior of those who share that name. This prestige and perceived honor are hence a function of how the holders of the name behave with others, whether they act fairly or not in general. Therefore, if we can account for the use of names in interactions, then we may account for the ones that are the most beneficial or the most honorable, and, as a consequence, for the ones that would preferably given to children as a function of the existing name distribution. Giving the name of someone one wants to honor to a child could also be accounted for if more people bearing the name in question within the family increases some measure of honorability for that name.

In order to build a model appropriate to the context of ancient Greece, we combine features of several of the economic approaches we have mentioned earlier. We consider groups of people, which we call families, in the sense of clans or clienteles, rather than a small-size household. In our logic there should be a reasonably limited number of such families in a *polis*. Along with Carr and Landa (1983), we can assume that members of a same family need to behave well with each other, as they would be identified and punished otherwise, while when dealing with people who are not family members there would be more leeway for cheating. The interactions in which one might exchange names are those defined as *personalistic exchanges* by Carr and Landa (1983).

Although we do not need to specifically define identity (as a function of the name or of the family), some of the results from Akerlof and Kranton (2000) may be derived at equilibrium due to optimal behavior of the whole population: people may behave in a certain way with those with whom they share the same name, which is an identity-charged feature. Finally, the notion of conformity as defined and exploited in Bernheim (1994) does not either required to be formally expressed in our approach, but we will see that a conformist behavior for name giving rises up in certain conditions.

2 Modeling the Use of Names

We assume there is a continuous population of individuals $x \in \mathcal{P}$, and to each individual is associated a single family $\phi_x \in \mathcal{F}$, a finite set of cardinal I. We will consider a probability measure μ over $\mathcal{P}$ that gives the relative frequency of families. We will write Φ_f for the set of individuals with family f, so that $\mu(\Phi_f)$ is properly defined. We further assume that for any element $x \in \mathcal{P}$, $\mu(\{x\}) = 0$.

In our modeling, we will use independent random variables X and Y following a uniform distribution on $\mathcal{P}$, and define a probability $\mathbb{P}$, so that for all μ-measurable set $A \subset \mathcal{P}$, $\mathbb{P}[X \in A] = \mu(A)$. One simple way of representing the families is with a series of positive values $(p_{f_i})_{1 \le i \le I}$ such that $\sum_i p_{f_i} = 1$ and for each element $f_i \in \mathcal{F}$, $\mu\left(\Phi_{f_i}\right) = p_{f_i}$.

All the people interact in pairs in the course of their life, and when they do, they can choose to act fairly, or to cheat. Incentives are in place to try and deter people from acting unfairly within the family: when they do, they receive a punishment (some reduction in utility, or some reparation that needs to be paid). When someone acts unfairly with respect to a member of their own family, they will be found individually. When people act unfairly with others outside of the family, they cannot be easily found and as a result no cost is directly assessed. Although there may be laws that punish crimes in general, the cheating we are describing here should not be a serious offence. As such, we consider that in practice some laws or rules that address the cheating could be enforced within a family, and much less so outside of a family.

When an individual $x \in \mathcal{P}$ interacts with $y \in \mathcal{P}$ and neither cheats, both gain g. If x cheats and y does not, then x gains $g + r$, but there is a cost: if x happens to be from the same family as y, ie $\phi_x = \phi_y$ then the cost is c_w. When both players cheat at the same time they have no gain and no loss. When a player does not cheat, but is cheated, there is a loss of r. We assume that $c_w > r$, so that the punishment always takes away at least as much as the ill-gotten gains. One could think of the punishment as a multiple of the potential gain, because such cheating reflects badly on the entire family. This setting is very simple, but refinements such as different costs schedules attached to participating, or when cheating outside of the family, do not change the overall shape of the results while at the same time obscuring the logic.

We will write the expectation operator $\mathbb{E}$ with respect to $\mathbb{P}$, and use μ to denote the relative size of population groups. In particular, note that $\mathbb{P}[\phi_X = \phi_Y] = \mathbb{E}\left[\mu\left(\Phi_{\phi_X}\right)\right]$. If we write the event $H_{x,y} \in \{0, 1\}$ when player x cheats on player y, then the gain for x in an interaction with y writes:

$$G_{x,y} = g\,\overline{H_{x,y}}\,\overline{H_{y,x}} + \left(g + r - c_w \mathbb{I}_{\phi_x = \phi_y}\right) H_{x,y}\overline{H_{y,x}} - r\,\overline{H_{x,y}}\,H_{y,x}.$$

The choice that an individual x can make is to cheat or not, which we will write H_x, conditioning on the information they may have (on their own family and name, or on the other's family and name). The core setting of this problem is comparable to a prisonner's dilemma, where cheating corresponds to defecting, and not cheating to cooperation. Considering mixed strategies, the optimal probability that individual x cheats with individual y will be written $\eta_{x,y}$, and the expected gain for individual x will be denoted by π_x. Both measures depend on the information available to each individual. We will write F_x the information set that individual x can condition upon, with random variables X and Y representing the randomly selected individuals in an encounter.

From a general perspective, we can write the expected gain for player x conditioned by the fact that x is cheating or not, and conditioned by x's available information:

$$\pi_x(H) = \mathbb{E}\left[G_{X,Y} \mid \{H_{X,Y} = 1\} \cap F_x\right]$$
$$\pi_x\left(\overline{H}\right) = \mathbb{E}\left[G_{X,Y} \mid \{H_{X,Y} = 0\} \cap F_x\right].$$

At the optimum, when all participants play the mixed strategy, we will have for all $x \in \mathcal{P}$:

$$\mathbb{E}\left[G_{X,Y} \mid \{H_{X,Y} = 1\} \cap F_x\right] = \mathbb{E}\left[G_{X,Y} \mid \{H_{X,Y} = 0\} \cap F_x\right].$$

As the most realistic setting, in line with what we can understand of relationships in ancient Greece, we will eventually assume that when people meet, they disclose their names but they do not disclose what family they belong to, because either they do not wish to, or they may not be able to prove what they would state. We consider that the name itself is necessarily given truthfully, the potential decision to cheat coming only after one knows the other's name. As a result, there is no perfect knowledge of a counterpart's family. Information on the distribution of names in families and on their sizes is known to everyone, on the other hand.

We begin by characterizing several possible equilibria, depending on the information available to the population: successively perfect information, no information (not even on one's family), knowledge of one's family, and finally knowledge of one's family and of the other's name, which is the most representative case. Considering these various possible conditions will allow us to also consider the aggregate social optimum reached in each case.

2.1 Equilibria with Perfect Information

In the simplest case with full information, which we will denote with the superscript a, $F_x = F_y = \{\phi_X = \phi_x, \phi_Y = \phi_y\}$. We have the following result.

Proposition 2.1 (Interactions with Perfect Information)**.** *When individual x encounters individual y with perfect information, the probability that x cheats is 1 if x and y are from a different* *family, and* $1 - \frac{r}{c_w}$ *otherwise. The aggregate probability of cheating is*

$$\mathbb{E}[\eta^a_{X,Y}(\phi_X, \phi_Y)] = 1 - \frac{r}{c_w} \sum_{1 \leq i \leq I} p^2_{f_i},$$

and the aggregate gain is:

$$\mathbb{E}[\pi^a_X(\phi_X, \phi_Y)] = \frac{r}{c_w}(g + r - c_w) \sum_{1 \leq i \leq I} p^2_{f_i}.$$

Proof. The optimal probability of cheating depends on the information known by both, so we simply write:

$$\eta^a_{x,y}(\phi_x, \phi_y) = \mathbb{I}_{\phi_x = \phi_y} \eta^a_{x,y}(\phi_x, \phi_y)|_{\phi_x = \phi_y} + \mathbb{I}_{\phi_x \neq \phi_y} \eta^a_{x,y}(\phi_x, \phi_y)|_{\phi_x \neq \phi_y}.$$

The following normal form representation shows the game when the two participants are from different families. One can see that in this case, there is a pure strategy Nash equilibrium, when both cheat, that is $\eta^a_{x,y}(\phi_x, \phi_y)|_{\phi_x \neq \phi_y} = 1$, and the expected gain is null, $\pi^a_x(\phi_x, \phi_y)|_{\phi_x \neq \phi_y} = 0$.

Player y, $\phi_x \neq \phi_y$

		H	$\overline{H}$
Player x	H	$(0, 0)$	$(g + r, -r)$
	$\overline{H}$	$(-r, g + r)$	(g, g)

When people are from the same family, then the interaction can be represented as follows, and there are then two possible Nash equilibria in pure strategies: both cheat, or both cooperate.

Player y, $\phi_x = \phi_y$

		H	$\overline{H}$
Player x	H	$(0, 0)$	$(g + r - c_w, -r)$
	$\overline{H}$	$(-r, g + r - c_w)$	(g, g)

Solving for the mixed strategy optimum in this case, we obtain:

$$\eta^a_{x,y}(\phi_x, \phi_y)|_{\phi_x = \phi_y} = \frac{c_w - r}{c_w} = 1 - \frac{r}{c_w},$$

which is between 0 and 1 with our assumptions.

With this mixed strategy, the expected gain for any player x writes:

$$\pi_x^a(\phi_x, \phi_y)|_{\phi_x=\phi_y} = \frac{r}{c_w}(g - c_w + r),$$

which is positive if $g > c_w - r$.

In these conditions, the overall expected gain from these interactions can be written:

$$\mathbb{E}[\pi_X^a(\phi_X, \phi_Y)] = \frac{r}{c_w}(g + r - c_w)\mathbb{E}\left[\mu\left(\Phi_{\phi_X}\right)\right].$$

The aggregate probability of cheating is

$$\begin{aligned}\mathbb{E}[\eta_{X,Y}^a(\phi_X, \phi_Y)] &= \frac{c_w - r}{c_w}\mathbb{E}\left[\mu\left(\Phi_{\phi_X}\right)\right] + \left(1 - \mathbb{E}\left[\mu\left(\Phi_{\phi_X}\right)\right]\right) \\ &= 1 - \mathbb{E}\left[\mu\left(\Phi_{\phi_X}\right)\right]\frac{r}{c_w} \\ &= 1 - \frac{r}{c_w}\sum_{1\le i\le I} p_{f_i}^2,\end{aligned}$$

where we use the fact that: $\mathbb{E}\left[\mu\left(\Phi_{\phi_X}\right)\right] = \sum_{1\le i\le I} p_{f_i}^2$. □

When people from different families meet, the interaction is a straight prisonner's dilemma: the gain from defecting (cheating) is strictly superior to the gain when cooperating. When people from the same family interact, since $c_w > r$, this is not a typical prisonner's dilemma according to that definition, the gain from cheating is less than the gain attached to cooperation. Relying on a large set of empirical studies on prisonner's dillemas, Mengel (2018) and Gaechter, Lee, and Sefton (2020) have shown that on non-repeated games, cooperation probability is still potentially high, typically in the 20%-60% range, in spite of the optimality of defection. It is however largely driven by the degree of temptation (defined in these papers) which in our case is equal to $\frac{r}{g}$. Assuming that r may be as large as g, this temptation would drive the likelihood of cooperation to very low levels, which would in turn justify the punishment c_w, so that at least in some cases there is high likelihood of cooperation.

With the cost c_w, in the case of interactions within the same family, we can see that as the punishment cost increases, the probability of cheating increases. This may seem paradoxical, but if there is a very large cost of cheating then the more likely choice is for everyone to cheat so that the probability of anyone cheating while the other is not cheating becomes very small. If the punishment cost can be optimally chosen, it should be only slightly larger than the extra gain obtained when cheating a non-cheating counterparty. Systematically cheating is a way of ensuring nobody gets caught and punished.

The aggregate outcome from these games with full information does not perform well socially, and, as we will see, can be improved when information on one's family is kept private.

2.2 Equilibria with No Information

At the other extreme, if individuals do not know, or do not factor in as part of their choices, either their own family nor their counterpart's family, the information set verifies $F_x = \varnothing$. We will denote this situation with the superscript $\varnothing$. The optimal strategy for x, $\eta_x^{\varnothing}$, does not depend on any input and is in fact constant as a function of x.

Proposition 2.2 (Interactions with No Information)**.** *When individual x encounters individual y with no information, the probability that x cheats is constant and equal to:*

$$\eta_x^{\varnothing} = 1 - \frac{r}{c_w \left(\frac{r}{c_w} \vee \sum_{1 \le i \le I} p_{f_i}^2 \right)}.$$

The expected aggregate gain is:

$$\pi_x^{\varnothing} = g - (g + r) \left(1 - \frac{r}{c_w \left(\frac{r}{c_w} \vee \sum_{1 \le i \le I} p_{f_i}^2 \right)} \right).$$

Proof. In the expression for player x's expected gain depending on their strategy, the only conditioning is that by $H_{X,Y} = 1$ and hence:

$$\pi_x^{\varnothing}(H) = \mathbb{E}\left[(1 - \eta_Y^{\varnothing})(g + r - c_w \mathbb{I}_{\phi_Y = \phi_X}) \right],$$

and

$$\pi_x^{\varnothing}\left(\overline{H}\right) = \mathbb{E}\left[-r\eta_Y^{\varnothing} + g(1 - \eta_Y^{\varnothing}) \right].$$

Expressing the expected payoffs of this game in normal form, we can see that when $c_w \mathbb{E}\left[\mu\left(\Phi_{\phi_Y}\right)\right] > r$, the dominant strategy is for all players to cooperate.

Player y

Player x		H	$\overline{H}$
	H	$(0, 0)$	$(g + r - c_w \mathbb{E}\left[\mu\left(\Phi_{\phi_Y}\right)\right], -r)$
	$\overline{H}$	$(-r, g + r - c_w \mathbb{E}\left[\mu\left(\Phi_{\phi_Y}\right)\right])$	(g, g)

Equating $\pi_x^{\varnothing}(H)$ and $\pi_x^{\varnothing}\left(\overline{H}\right)$ in the cases where there is no dominant strategy, and solving for $\eta_Y^{\varnothing}$ (which does not depend on Y), we obtain:

$$\eta_x^{\varnothing} = 1 - \frac{r}{c_w\left(\frac{r}{c_w} \vee \mathbb{E}\left[\mu\left(\Phi_{\phi_Y}\right)\right]\right)}.$$

Since $\mathbb{E}\left[\mu\left(\Phi_{\phi_Y}\right)\right] = \sum_{1\leq i\leq I} p_{f_i}^2$, we can also write:

$$\eta_x^{\varnothing} = 1 - \frac{r}{c_w\left(\frac{r}{c_w} \vee \sum_{1\leq i\leq I} p_{f_i}^2\right)}.$$

The expected profit writes

$$\begin{aligned}
\pi_x^{\varnothing} &= -r\eta_Y^{\varnothing} + g(1-\eta_Y^{\varnothing}) \\
&= g - (g+r)\left(1 - \frac{r}{c_w\left(\frac{r}{c_w} \vee \sum_{1\leq i\leq I} p_{f_i}^2\right)}\right).
\end{aligned}$$

We can see how the aggregate probability of cheating enters the expression for the expected aggregate gain. □

If $c_w < \frac{r}{\mathbb{E}[\mu(\Phi_{\phi_Y})]}$, then the probability of encountering someone from the same family is low enough that the potentially large cost of cheating a non-cheating counterparty is reduced and does not require to be eliminated by everyone choosing to cheat at the same time, and the optimal strategy is for all to cooperate. If $c_w > rI$, it implies the probability of cheating is strictly positive.

It is easy to see that $\pi_x^a(\phi_x, \phi_y) \leq \pi_x^{\varnothing}$, since $\sum_{1\leq i\leq I} p_{f_i}^2 \leq 1$, and this situation with no information is more socially beneficial than when people have full knowledge.

2.3 Equilibria with Private Information

We now consider the case when in these interactions each person knows their own family, but not the other's. The information set for x is therefore $F_x = \{\phi_X = \phi_x\}$, and we will denote this case with the superscript o. We will look for a mixed strategy optimum, expressed as a probability $\eta_x^o(\phi_x)$ of cheating, as a function of the participant x's own family ϕ_x. The dependency of $\eta_x^o(\phi_x)$ to x is only through ϕ_x. Note that in this case, the distributions from which X and Y are drawn are the same.

Proposition 2.3 (Interactions with Private Information)**.** *When individual x encounters individual y with private information, the probability that x cheats as function of their information* *set is:*

$$\eta_x^o(\phi_x) = 1 - \frac{r}{c_w\left(\frac{r}{c_w} \vee \mu\left(\Phi_{\phi_x}\right)\right)}.$$

The aggregate probability of cheating across all possible interactions is:

$$\mathbb{E}[\eta_Y^o(\phi_Y)] = 1 - \frac{r}{c_w}\sum_{1\leq i\leq I}\frac{p_{f_i}}{\frac{r}{c_w} \vee p_{f_i}},$$

and the aggregate gain is:

$$\mathbb{E}\left[\pi_X^o\left(\phi_X, \phi_Y\right)\right] = g - (r+g)\left(1 - \frac{r}{c_w}\sum_{1\leq i\leq I}\frac{p_{f_i}}{\frac{r}{c_w} \vee p_{f_i}}\right).$$

Proof. The expected gain when cheating, and when the other participant y plays the optimal strategy $\eta_y^o(\phi_y)$ can be written:

$$\pi_x^o(H) = \mathbb{E}\left[(1-\eta_Y^o(\phi_Y))(g + r - c_w \mathbb{I}_{\phi_Y=\phi_X}) \mid \phi_X = \phi_x\right],$$

and the expected gain when not cheating is

$$\begin{aligned}\pi_x^o\left(\overline{H}\right) &= \mathbb{E}\left[-r\eta_Y^o(\phi_Y) + g(1-\eta_Y^o(\phi_Y)) \mid \phi_X = \phi_x\right] \\ &= \mathbb{E}\left[-r\eta_Y^o(\phi_Y) + g(1-\eta_Y^o(\phi_Y))\right].\end{aligned}$$

At the optimum, both strategies must have the same expected return for x, so that $\pi_x^o\left(\overline{H}\right) = \pi_x^o(H)$. Noting that

$$\mathbb{E}\left[\eta_Y^o(\phi_Y)\mathbb{I}_{\phi_Y=\phi_x}\right] = \eta_x^o(\phi_x)\mu\left(\Phi_{\phi_x}\right),$$

we obtain the equation:

$$r = c_w \mu\left(\Phi_{\phi_x}\right)(1 - \eta_x^o(\phi_x)),$$

and therefore

$$\eta_x^o(\phi_x) = 1 - \frac{r}{c_w\left(\frac{r}{c_w} \vee \mu\left(\Phi_{\phi_x}\right)\right)}.$$

If there are very few people in x's family, then there is a low probability of encountering people from the same family and needing to systematically cheat in order to optimally avoid the large punishment, and it is optimal to cooperate. Reciprocally, the more people there are in the family, the greater the cheating probability.

We can compute the average optimal probability $\mathbb{E}[\eta_{\phi_Y}]$ by integrating:

$$\begin{aligned}\mathbb{E}[\eta_Y^o(\phi_Y)] &= 1 - \mathbb{E}\left[\frac{r}{c_w\left(\frac{r}{c_w} \vee \mu\left(\Phi_{\phi_Y}\right)\right)}\right] \\ &= 1 - \frac{r}{c_w}\sum_{1\le i\le I}\frac{p_{f_i}}{\frac{r}{c_w} \vee p_{f_i}}.\end{aligned}$$

We can compute the expected gain across individuals:

$$\begin{aligned}\mathbb{E}\left[\pi_X^o\left(\phi_X, \phi_Y\right)\right] &= \mathbb{E}[-r\eta_Y^o(\phi_Y) + g(1-\eta_Y^o(\phi_Y))] \\ &= g - (r+g)\left(1 - \frac{r}{c_w}\sum_{1\le i\le I}\frac{p_{f_i}}{\frac{r}{c_w} \vee p_{f_i}}\right).\end{aligned}$$

Here again, we can directly see how the expected probability of cheating enters the aggregate gain expression. □

We can recognize a form similar to the one for the gain when all information is known and the players are from the same family.

The difference between the expected probability of cheating with private information and with no information at all writes:

$$\mathbb{E}[\eta_Y^o(\phi_Y)] - \mathbb{E}[\eta_Y^\varnothing(\phi_Y)] = \frac{r}{c_w\left(\frac{r}{c_w} \vee \sum_{1\le i\le I} p_{f_i}^2\right)} - \frac{r}{c_w}\sum_{1\le i\le I}\frac{p_{f_i}}{\frac{r}{c_w} \vee p_{f_i}}.$$

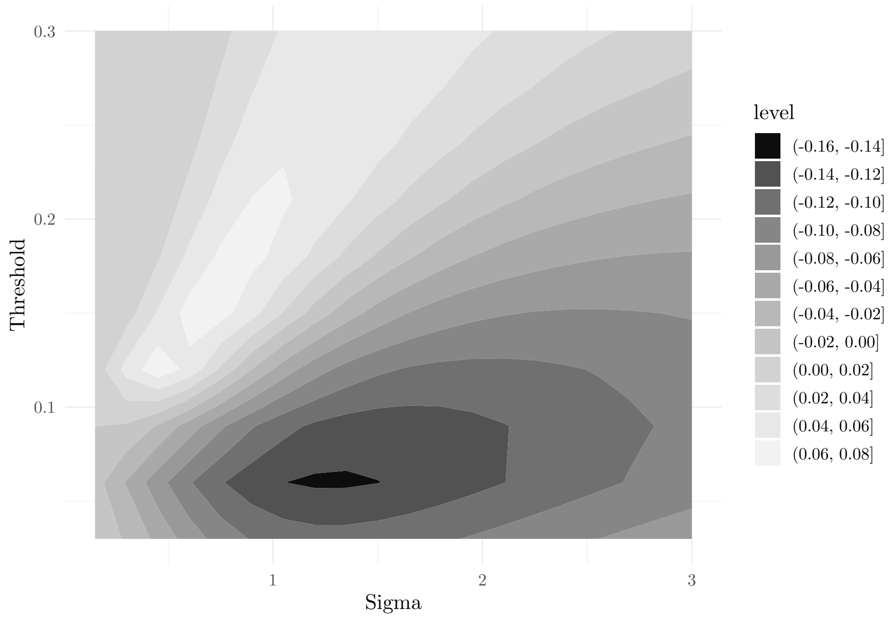
If the family size is regularly spread out and $p_{f_i} = \frac{1}{I}$, then

$$\mathbb{E}[\eta_Y^o(\phi_Y)] = \mathbb{E}[\eta_Y^\varnothing(\phi_Y)] = 1 - \frac{r}{c_w}\left(I \vee \frac{c_w}{r}\right).$$

Figure 1 shows the difference in the cheating probabilities between the private information case and the case with no information at all across a large range of simulated population distributions. The figure illustrates the fact that there is no systematic relationship between the two probabilities. When the threshold $\frac{r}{c_w}$ is high, the cheating probability with private information appears generally higher.

However, if $\frac{r}{c_w}$ is small and $\frac{r}{c_w} < \frac{1}{I}$, then $\frac{r}{c_w\left(\frac{r}{c_w} \vee \sum_{1\le i\le I} p_{f_i}^2\right)} = \frac{r}{c_w \sum_{1\le i\le I} p_{f_i}^2}$. In addition, if for all f, $p_f \geq \frac{r}{c_w}$, then $\sum_{1\le i\le I}\frac{p_{f_i}}{\frac{r}{c_w} \vee p_{f_i}} = I$ and $\mathbb{E}[\eta_Y^o(\phi_Y)] = 1 - \frac{r}{c_w}I$. In these conditions, if $p_f \neq \frac{1}{I}$, then $\frac{r}{c_w \sum_{1\le i\le I} p_{f_i}^2} < I\frac{r}{c_w}$, and $\mathbb{E}[\eta_Y^o(\phi_Y)] < \mathbb{E}[\eta_Y^\varnothing(\phi_Y)]$. When the punishment for cheating is costly relative to its potential gains, then having access to private information reduces

Figure 1: *Difference in Cheating Probability with Private Information and with No Information*



Note: The count of each family is drawn randomly from a lognormal distribution with dispersion parameter σ. The numbers shown in the chart are the expected values across a large number of simulations. The threshold is $\frac{r}{c_w}$ and $I = 10$.

the probability of cheating. In effect, if the distribution of people across families is spread out enough relative to $\frac{r}{c_w}$, then having private information is not detrimental to the social good.

2.4 Equilibria With Partially Common Information

We now address a generalization of the equilibrium with private information, considering that the distributions from which the participants are drawn are not the same. We assume that for each $x \in \mathcal{P}$, there exists μ_x, a measure giving the relative weight of the population sharing the same characteristics as x, and associated probability $\mathbb{P}_x$ such that for all measurable $A \subset \mathcal{P}$, $\mathbb{P}_x[X \in A] = \mu_x(A)$. All the μ_x represent the distribution of the population conditioned by some characteristic specific to x, for example x's name. Naturally, we may have $\mu_x = \mu_z$ for some x and z in $\mathcal{P}$, for example two individuals with the same name.

We consider the case of interactions when individuals x and y know their own family, but not the other person's, and the measures μ_x and μ_y are known by both. This situation is denoted with the superscript n. The optimal cheating probability can be written $\eta_x^n(\phi_x, \mu_y)$, and it depends on the knowledge of x's family ϕ_x and on the distributional information on y's characteristics. In this context, μ_x and μ_y are given, just as μ was a given in the prior cases. The expression $\mu_y\left(\Phi_{\phi_x}\right)$ captures the probability that x would be in the same family as y knowing x's characteristics, based on the knowledge of y's family.

If the μ_x represent names, then these μ_x can be represented by the mass of individuals with a given name in a given family. Set $\nu_x \in \mathcal{N}$, with $\mathcal{N}$ finite, the name of individual x. We also define $p_n = \mu(\{x \in \mathcal{P} : \nu_x = n\}) = \sum_{f \in \mathcal{F}} p_{\nu_x, f}$ and as before, $p_f = \mu(\{x \in \mathcal{P} : \phi_x = f\}) = \mu(\Phi_f)$. Then $\mu_x(\Phi_f) = \frac{p_{\nu_x, f}}{p_{\nu_x}}$. All the $p_{n,f}$ can naturally be expressed as a function of the population ρ in a given family with a given name: $p_{n,f} = \frac{\rho_{n,f}}{\sum_{m \in \mathcal{N}, g \in \mathcal{F}} \rho_{m,g}}$.

Proposition 2.4 (Interactions with Partially Common Information)**.** *When individual x encounters individual y with partially common information, the probability that x cheats as function of their information set (which contains their own family and distributional information on y) is:*

$$\eta_x^n(\phi_x, \mu_y) = 1 - \frac{r}{c_w \left(\frac{r}{c_w} \vee \mu_y(\Phi_{\phi_x}) \right)} = 1 - \left(1 \wedge \frac{r}{c_w} \frac{1}{\mu_y(\Phi_{\phi_x})} \right).$$

In aggregate, the cheating probability across all possible encounters is:

$$\mathbb{E}[\eta_X^n(\phi_X, \mu_Y)] = 1 - \frac{r}{c_w} \sum_{n \in \mathcal{N}} p_n \sum_{f \in \mathcal{F}} \frac{p_f}{\frac{r}{c_w} \vee \frac{p_{n,f}}{p_n}}.$$

Proof. The expressions for x's expected gain writes as follows:

$$\begin{aligned} \pi_x^n(H) &= \mathbb{E}_y \left[(1 - \eta_Y^n(\phi_Y, \mu_x))(g + r - c_w \mathbb{I}_{\phi_Y = \phi_x}) \right] \\ \pi_x^n\left(\overline{H}\right) &= \mathbb{E}_y \left[-r \eta_Y^n(\phi_Y, \mu_x) + g(1 - \eta_Y^n(\phi_Y, \mu_x)) \right]. \end{aligned}$$

Swapping x and y, and equating the expected returns $\pi_x^n(H)$ and $\pi_x^n\left(\overline{H}\right)$, we get:

$$\eta_x^n(\phi_x, \mu_y) = 1 - \frac{r}{c_w \left(\frac{r}{c_w} \vee \mu_y(\Phi_{\phi_x}) \right)}.$$

The probability that x cheats y is therefore increasing as a function of the chances that y would be in the same family as x conditioned on x's name; this is a comparable effect to what we observed earlier in the probability of cheating.

The expected gain for x hence writes:

$$\mathbb{E}[\pi_x^n] = g - (g + r) \left(1 - \frac{r}{c_w} \mathbb{E} \left[\frac{1}{\frac{r}{c_w} \vee \mu_Y(\Phi_{\phi_x})} \right] \right).$$

The probability of cheating can be integrated relative to $\mathbb{P}_x$ and $\mathbb{P}_y$ to get the aggregate probability of cheating.

$$\mathbb{E}[\eta_X^n(\phi_X, \mu_Y)] = 1 - \mathbb{E}\left[\frac{r}{c_w\left(\frac{r}{c_w} \vee \mu_Y\left(\Phi_{\phi_X}\right)\right)}\right].$$

We can write for any measurable function h:

$$\begin{aligned}\mathbb{E}\left[h\left(\mu_Y\left(\Phi_{\phi_X}\right)\right)\right] &= \int h\left(\mu_y\left(\Phi_f\right)\right)\mathbb{P}[Y \in dy]\mathbb{P}[f \in df] \\ &= \sum_{f\in\mathcal{F}} p_f \sum_{n\in\mathcal{N}} p_n h\left(\frac{p_{n,f}}{p_n}\right).\end{aligned}$$

Hence we have:

$$\mathbb{E}\left[\frac{r}{c_w\left(\frac{r}{c_w} \vee \mu_Y\left(\Phi_{\phi_X}\right)\right)}\right] = \frac{r}{c_w}\sum_{n\in\mathcal{N}} p_n \sum_{f\in\mathcal{F}} \frac{p_f}{\frac{r}{c_w} \vee \frac{p_{n,f}}{p_n}},$$

and

$$\mathbb{E}[\eta_X^n(\phi_X, \mu_Y)] = 1 - \frac{r}{c_w}\sum_{n\in\mathcal{N}} p_n \sum_{f\in\mathcal{F}} \frac{p_f}{\frac{r}{c_w} \vee \frac{p_{n,f}}{p_n}}$$

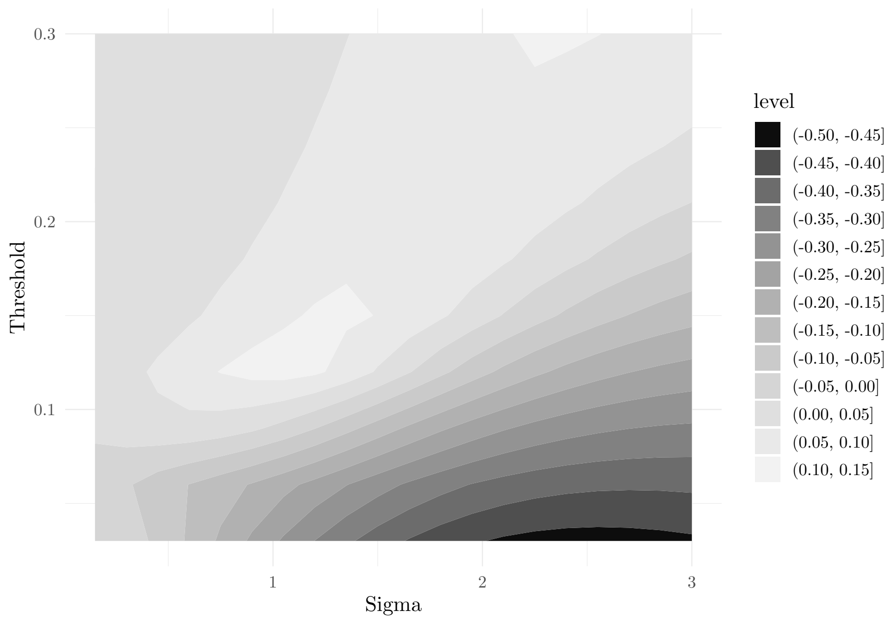
where we can recognize a form similar to that of $\mathbb{E}[\eta_Y^o(\phi_Y)]$. □

Figure 2 shows the difference in the cheating probabilities between the partially private information case and the case with no information, and Figure 3 shows the difference with the private information case. While there appear to be some patterns, there is no systematic relationships between these quantities and no systematically better system. When the threshold $\frac{r}{c_w}$ is low (high cost of being caught), it appears based on Figure 2 that the cheating probability with partially common information is lower than with no information. There is a similar effect at play in Figure 3.

The expression for $\mathbb{E}[\eta_X^n(\phi_X)]$ shows that if and only if, for all f and n, $\frac{p_{n,f}}{p_n} > \frac{r}{c_w}$, then $\mathbb{E}[\eta_X^n(\phi_X)] = 1 - \frac{r}{c_w}I$. We know that if and only if for all f, $p_f > \frac{r}{c_w}$, then $\mathbb{E}[\eta_X^o(\phi_X)] = 1 - \frac{r}{c_w}I$. The first condition implies the second one, and as a result, if the second condition is verified, $\mathbb{E}[\eta_X^o(\phi_X)] \leq \mathbb{E}[\eta_X^n(\phi_X)]$.

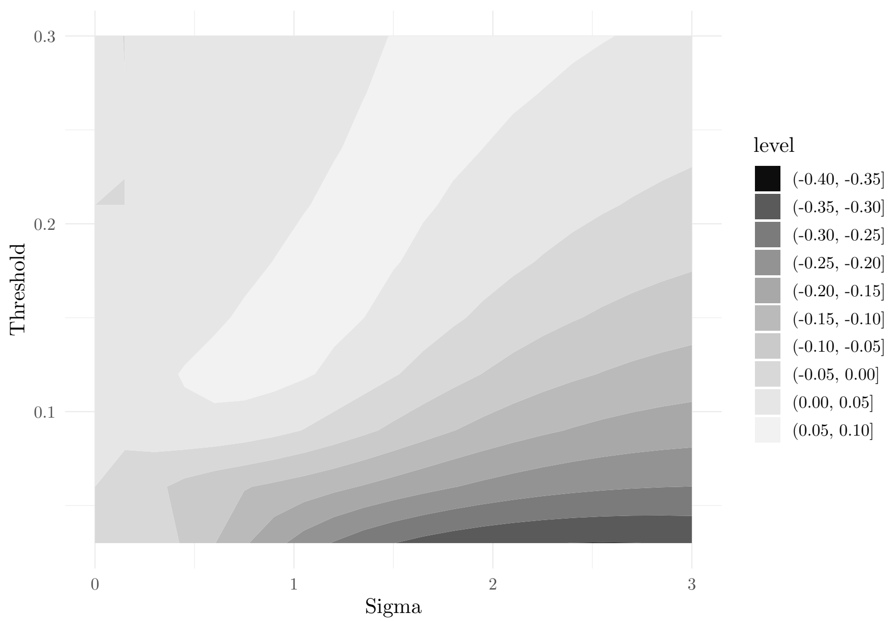
We can see that if the distribution of names is well spread out across families, then this knowledge does not condition much, and as a result the aggregate gain is comparable to the one with only private information or the one with no information at all. In particular, if the punishment c_w is severe enough, then the conditional probabilities are likely greater than $\frac{r}{c_w}$ and the aggregate social gain is the same as if there was no information. The use of names as important information, consistent with the way we represent interactions in ancient Greece, does not lead to a lower social optimum in this case than in the no information case. The exchange of names can hence fulfill a

Figure 2: *Difference between the Cheating Probability with Partially Common Information and with No Information*



Note: The count of each name in each family is drawn randomly from a lognormal distribution with dispersion parameter σ. The numbers shown in the chart are the expected values across a large number of simulations. The threshold is $\frac{r}{c_w}$, $|\mathcal{N}| = 12$ and $I = 10$.

Figure 3: *Difference between the Cheating Probability with Partially Common Information and with Private Information*



Note: The count of each name in each family is drawn randomly from a lognormal distribution with dispersion parameter σ. The numbers shown in the chart are the expected values across a large number of simulations. The threshold is $\frac{r}{c_w}$, $|\mathcal{N}| = 12$ and $I = 10$.

social role, by allowing people to give each other something before entering in a more demanding relationship, which would not be necessary in the three alternate information frameworks we discussed earlier, without a significant social cost relative to these alternatives.

We can see that the names which would appear as the most attractive according to this logic are the ones that are most common outside of the family: one wants to maximize $\frac{\rho_{a_j}}{\rho_{a_j,f}}$.

3 The Dynamics of Naming Choices

We have concentrated on understanding a possible function of names, and on some reasons why names could be used in interpersonal exchanges. These names, being given to the children by the parents, had to be determined in some way, and we now examine the way in which they could be chosen within the framework of our simple model. We will consider several drivers of name choices, the strategic one based on the probability of cheating being just one of them.

In light of the historical context we mentioned earlier, we consider that the choice of the names observed in a Greek *polis* may come from three distinct sources: a strategic choice, traditional paponymy, or the random appearance of new names. We will consider that the stock of each name grows through time as a function of these three sources. Since the choice of names is applicable to newborns, it affects a fraction of the population being born over each instant dt. We consider that children belong to a unique family, which could be that of the mother or the father. Individuals will pick the name with the greatest attractiveness, or of a parent, or from some external source at any point, without having to factor in the behavior of the others, since the change in population is assumed to be very small relative to the total population size.

In order to represent paponymy within the framework of our model, we will transform it into a simple patronymy, so that to a certain extent, names simply grow in the population at a rate proportional to their current stock. In a continuous time setting, the alternating jump over a generation should not matter. In addition, as we will see, the data does not allow for a detailed prosopographic perspective (tracking each life history), but only an onomastic perspective (tracking the names).

3.1 Representing Strategic Name Choices

In the interactions between people according to the model we have developped so far, we have seen that the individual's characteristics affect both the expected gain, which depends on the probability of one being cheated, and the probability of one cheating on others. The greatest the probability that others cheat, the smallest the gain. Hence, the individuals for whom the characteristics μ_x are such that the expected gain is highest are better off. Varying the characteristics of x and y, the probability that x cheats writes $\eta_x^n(\phi_x, \mu_y)$, and depends on x's family and on y's name, written as a function of the population distribution ρ:

$$\eta_x^n\left(\phi_x, \mu_y\right) = 1 - 1 \wedge \frac{r}{c_w} \frac{\rho_{\nu_y}}{\rho_{\nu_y, \phi_x}}.$$

We can consider the desirability of a name to be related to the probability of not cheating, that is the probability that others do not cheat on a given person. The lowest the probability, the higher the expected outcome, and the better off the person. Also, from an honor standpoint, it should presumably be perceived as more honorable to not be cheated. We can note nevertheless that when c_w is near r, then the choice does not have much of an impact, because the cheating probability is low anyway. If c_w is large, on the other hand, then picking one name or the other could have a significant impact on the amount of cheating the individual does or suffers. The attractivity of a name a may be defined as the probability that people in general do not cheat on a, taking values in $[\frac{r}{c_w}, 1]$:

$$A(a, \rho) = \sum_f \frac{\rho_f}{\rho} \left(1 \wedge \frac{r}{c_w} \frac{\rho_a}{\rho_{a,f}}\right).$$

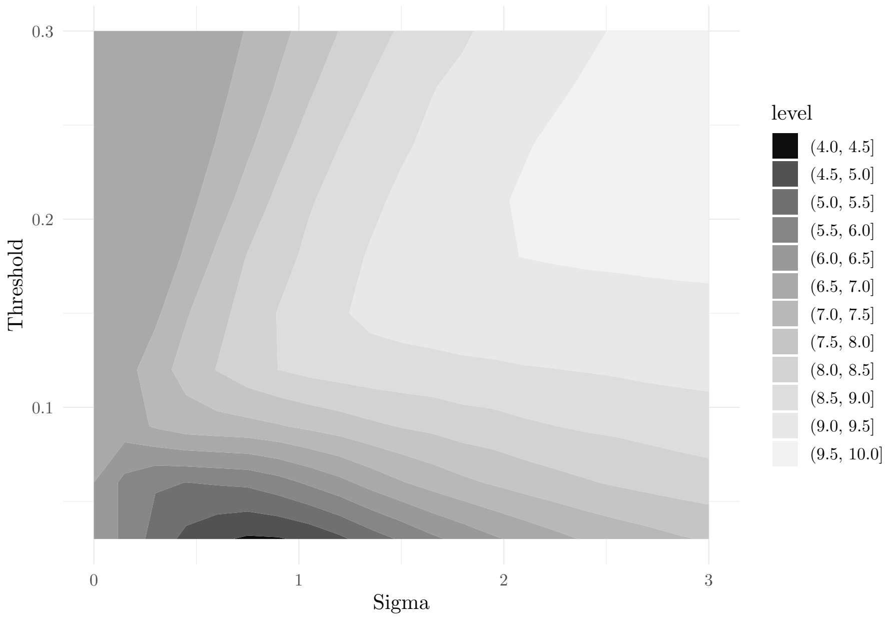
The optimal name choice depends on the manner in which all names and families are distributed. In studying the dynamics of name distributions, fully reflecting this complexity makes it impossible to obtain clear results. In order to illustrate some aspects of the optimal name choice, Figure 4 shows the average rank (in terms of popularity) of the name that gets picked across a large range of simulated population layouts, following the same logic as in the prior figures. We can see that there are some substantial variations in the nature of the names being picked, depending on each specific situation.

Proposition 3.1 (Optimal Name Choice Logic)**.** *If names are distributed in the same manner in each family, then the attractivity is constant. If names are concentrated in equally sized families, then:*

- *If the most common name is not very frequent, then it is optimal to chose the least common name;*
- *If the most common name is frequent, then it is optimal to chose the most common name.*

Proof. We will first consider the case when names are distributed in the same way across

Figure 4: *Rank by Number of Carriers of the Optimal Name that Minimizes the Probability of Being Cheated*



Note: The count of each name in each family is drawn randomly from a lognormal distribution with dispersion parameter σ. The numbers shown in the chart are the expected values across a large number of simulations. The threshold is $\frac{r}{c_w}$, $|\mathcal{N}| = 12$ and $I = 10$.

families. We assume that for all name a, $\frac{\rho_{a,f}}{\rho_f} = \rho_a$. Applying this simplification we get $A(a,\rho) = \sum_f p_f \left(1 \wedge \frac{r}{c_w p_f}\right)$, which does not depend on a.

If names are concentrated in families, this means we can write $p_{a,f} = p_f \mathbb{I}_{f \in F_a}$, where F_a is defined as the subset of families that carry name a. Also, for further simplification, we assume that the families are sized equally, and $p_f = \frac{1}{I}$. Hence, $p_a = \frac{|F_a|}{I}$. We can write:

$$\begin{aligned} A(a,\rho) &= \frac{1}{I}\sum_f \left(1 \wedge \frac{r}{c_w}\frac{Ip_a}{\mathbb{I}_{f\in F_a}}\right) \\ &= \frac{1}{I}\sum_f (1-\mathbb{I}_{f\in F_a}) + \frac{1}{I}\sum_f \mathbb{I}_{f\in F_a}\mathbb{I}_{Ip_a\frac{r}{c_w}>1} + \frac{1}{I}\sum_f \mathbb{I}_{f\in F_a}\frac{r}{c_w}p_a I \mathbb{I}_{Ip_a\frac{r}{c_w}<1} \\ &= 1 + \mathbb{I}_{Ip_a\frac{r}{c_w}<1}\left(p_a^2 I\frac{r}{c_w} - p_a\right). \end{aligned}$$

Hence, if $p_a > \frac{c_w}{2Ir}$, A is increasing as a function of p_a, and decreasing otherwise. Therefore, if the most frequent name's frequency is high enough it is optimal to pick it. If it is under the threshold, it is optimal to pick the least frequent name. If $p_a > \frac{c_w}{Ir}$ however, then the attractivity is maximal, and any name that satisfies this condition may be chosen. □

The strategy of picking new names could hence appear as either leaning towards innovative fashion, by choosing an initially rare and new name that is picked by everybody, or towards a conformist choice, of picking the most common existing names. At the limit in either case, the informational content from communicating the name would become low.

In the simplified context of the Proposition 3.1, if it is optimal to pick the rarest name, then if new names sometimes appear (being invented or through immigration), then these names would naturally seem optimal, because they would be very rare at first. Also, as the rarest names are picked, they become less rare until eventually they would catch up with other names. This situation could lead to a uniform distribution of names, where names would then not convey particular information. If on the contrary the most frequent name is picked, then it may become the only existing name, in which case the informational content in the names would also vanish.

More generally, we can see that if the cost c_w is not extreme and there are many families, then the optimal choice would appear to oscillate between the most frequent names, or just any name. Without being able to characterize the optimal choice more precisely, it is reasonnable to approximate it by a random choice weighed by name frequency, as a mix between a selection of the most common names and a totally random choice.

3.2 A Model for the Distribution of Names

In the ancient Greek population for a *polis*, we consider that the natural population net growth rate is $\lambda - \mu$, where λ is the birth or immigration rate, and μ is the death or emigration rate. When people appear (whether they are born or immigrants), there are three possible manners in which their name is determined:

- The name may be given as a strategic choice according to the model we developed earlier, in which case it will be selected using a particular logic. The rate at which this may happen per unit of time, as a function of the number of people, is γ;
- The name may be a new name that did not exist before. This could reflect the effect of immigration, or simply some creativity in naming. The rate at which new names may be acquired per unit of time, as a function of the number of people, is β;
- Otherwise, the name is given using paponymy, or in our context patronymy, at the rate of $\lambda - \beta - \gamma$. We assume that $\lambda > \beta + \gamma$ so that only newly born people get a non-patronimic name.

With this approach, we follow certain aspects of the model from Baek, Kiet, and Kim (2007), however in their model they considered that names were changed after people were born, and did not have strategic name choices. We keep our notations ρ for the count of people in the population, and N as the total number of names, but now depending on time since we are in a dynamic evolution context. The overall approach is also consistent with the model proposed by Reed and Hughes (2003), who also relied on transition probabilities and on probability generating functions.

Proposition 3.2 (Evolution of the Number of Names)**.** *The average number of names per people verifies*

$$\frac{N(t)}{\rho(t)} = \frac{\beta}{\lambda - \mu}\frac{e^{(\lambda-\mu)t} - 1}{e^{(\lambda-\mu)t}}$$

and converges to $\frac{\beta}{\lambda-\mu}$ *as* t *goes to infinity.*

Proof. The total population is not affected by how people are named, it only depends on the net growth rate α, and

$$\rho(t) = \rho(0)e^{(\lambda-\mu)t}.$$

We define $\pi(s)$ as the rate of names appearing at a point in time s, that is $\pi(s) = \beta\rho(s)$. Hence the total number of names at time t verifies $N(t) = \int_0^t \pi(s)ds$. As a result, we have:

$$\frac{N(t)}{\rho(t)} = \frac{\int_0^t \pi(s)ds}{\rho(0)e^{(\lambda-\mu)t}} = \frac{\int_0^t \beta e^{(\lambda-\mu)s}ds}{\rho(0)e^{(\lambda-\mu)t}} = \frac{\beta}{\lambda-\mu}\frac{e^{(\lambda-\mu)t} - 1}{e^{(\lambda-\mu)t}}.$$

Hence $\lim_{t\to\infty} \frac{N(t)}{\rho(t)} = \frac{\beta}{\lambda-\mu}$. □

We now define the conditional probability that there are k people carrying the name n at time t, given there were j people carrying it at time s:

$$p_n(j, k, s, t) = \mathbb{P}\left[\rho_n(t) = k \mid \rho_n(s) = j\right].$$

We can express the dynamics of the number of people carrying a given name over a short period of time dt as follows. We know that the part of births named by paponymy $\lambda - \beta - \gamma$ can explain an increase in the population carrying the name. The rate β, which refers to new names being created, can only take people away from existing names and does not add them to existing ones. The rate γ, concerning strategic naming, takes people away from a given name, but also redistributes them across all existing names. The overall dynamics for the population, in terms of the number of names or number of people, need to be consistent with the way transitions in each cohort's population are expressed[9].

Within this framework, a cohort with $k+1$ people could only transition to k if one of them died or emigrated, at the rate μ. Therefore, we shoud assume that over a short time period dt:

$$\mathbb{P}\left[\rho_n(t+dt) = k \mid \rho_n(t) = k+1\right] = (k+1)\mu dt.$$

A cohort with $k-1$ people could transition to k if one person was born and named according to paponymy, or one person was strategically named to the current cohort. The probability for this last event may depend on the distribution of all other names, so for now we will write it as $\gamma(\rho, p, k-1)$. Hence, we assume that:

$$\mathbb{P}\left[\rho_n(t+dt) = k \mid \rho_n(t) = k-1\right] = (k-1)(\lambda - \beta - \gamma)dt + \gamma(\rho, p, k-1)dt.$$

Finally, the probability that a cohort with k people would remain with the same number of people is the probability that none of the possible changes we discussed above takes place. Hence, we would expect that:

$$\mathbb{P}\left[\rho_n(t+dt) = k \mid \rho_n(t) = k\right] = 1 - k\left(\lambda + \mu - \beta - \gamma\right)dt - \gamma(\rho, p, k-1)dt.$$

Reflecting the strategic name choice, event with the simplest assumption such as picking a uniform distribution across names, makes the calculation of the limiting behavior impossible. In the case where there is no strategic choice, that is $\gamma = 0$, we can however derive some results following Baek, Kiet, and Kim (2007). Although even in this case our model setup is different from theirs (people are named at birth, do not change name afterwards), we find the same limit distribution.

Proposition 3.3 (Transition Probability Expression)**.** *For a small enough increment dt, the transition probability p_n verifies:*

$$\begin{aligned}\frac{dp_n(j,k,s,t)}{dt} &= -k\left(\lambda + \mu - \beta\right)p_n(j,k,s,t) \\ &+ (k-1)\left(\lambda - \beta\right)p_n(j,k-1,s,t) \\ &+ (k+1)\mu p_n(j,k+1,s,t).\end{aligned}$$

[9] The introduction of a non proportional increase in new names in Baek, Kiet, and Kim (2007) does not appear to be consistent with the transition dynamics they assume.

Proof. Considering an arbitrarily small time increment $dt > 0$, we have:

$$\begin{aligned}
p_n(j,k,s,t+dt) &= \mathbb{P}\left[\rho_n(t+dt) = k \mid \rho_n(s) = j\right] \\
&= \mathbb{E}\left[\mathbb{I}_{\rho_n(t+dt)=k} \mid \rho_n(s) = j\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[\mathbb{I}_{\rho_n(t+dt)=k} \mid \rho_n(t)\right] \mid \rho_n(s) = j\right] \\
&= \sum_l \mathbb{P}\left[\rho_n(t) = l \mid \rho_n(s) = j\right] \mathbb{P}\left[\rho_n(t+dt) = k \mid \rho_n(t) = l \cap \rho_n(s) = j\right].
\end{aligned}$$

However, since dt is small, we will assume that ρ_n cannot change by more than 1, that is, for all k and l:

$$\mathbb{I}_{|k-l|>1}\mathbb{P}\left[\rho_n(t+dt) = k \mid \rho_n(t) = l \cap \rho_n(s) = j\right] = 0.$$

Hence, we have

$$\begin{aligned}
p_n(j,k,s,t+dt) = p_n(j,k,s,t)&\mathbb{P}\left[\rho_n(t+dt) = k \mid \rho_n(t) = k \cap \rho_n(s) = j\right] \\
&+ p_n(j,k-1,s,t)\mathbb{P}\left[\rho_n(t+dt) = k \mid \rho_n(t) = k-1 \cap \rho_n(s) = j\right] \\
&+ p_n(j,k+1,s,t)\mathbb{P}\left[\rho_n(t+dt) = k \mid \rho_n(t) = k+1 \cap \rho_n(s) = j\right].
\end{aligned}$$

The particular expressions for the law of $\rho_n(t+dt)$ in our assumptions do not depend on $\rho_n(s)$. Replacing in the expression, we can write:

$$\begin{aligned}
p_n(j,k,s,t+dt) - p_n(j,k,s,t) &= -k\left(\lambda + \mu - \beta\right) p_n(j,k,s,t)dt \\
&+ (k-1)\left(\lambda - \beta\right) p_n(j,k-1,s,t)dt \\
&+ (k+1)\mu p_n(j,k+1,s,t)dt.
\end{aligned}$$

This gives us the statement in the proposition. □

Focusing on the case when $j = 1$, that is when we are conditioning by the point when there was only one person in the cohort, we want to calculate $p_n(1,k,s,t)$. In the case where $\gamma = 0$, this probability does not explicitly depend on n, as it is the same expression for all names.

3.3 Name Distribution Limiting Behavior

We are now interested in the behavior of the name distribution as t becomes large. We write $P_n(t,k)$ the probability that there are k people with name n at time t, without any conditioning. Also, we write A_n the time when a name n appeared. If we are not treating the names differently (as is the case when $\gamma = 0$, we do not need to keep the subscripts. We have

$\mathbb{P}[A \in ds | A \leq t] = \frac{\pi(s)}{\int_0^t \pi(s)ds}$, since π tracks the rate of occurrence of new names. Therefore, conditioning the transition probabilities by the time of apparition of the name, we can write

$$P(t,k) = \int_0^t p(s,t,k)\mathbb{P}[A \in ds | A \leq t] = \frac{\int_0^t p(s,t,k)\pi(s)ds}{\int_0^t \pi(s)ds}.$$

We have the following result, which, being identical to that from Baek, Kiet, and Kim (2007) shows that the differences in the model set up did not have an influence.

Proposition 3.4 (Limiting Distribution without Strategic Naming)**.** *If* $\gamma = 0$*, the limiting distribution when* t *goes to infinity* P_∞ *verifies* $P_\infty(k) \sim k^{2+\frac{\beta}{\lambda-\beta-\mu}}$.

Proof. Using the explicit form for π, we have:

$$P(t,k) = \frac{\lambda - \mu}{1 - e^{-(\lambda-\mu)t}} \int_0^t e^{-(\lambda-\mu)(t-s)} p(s,t,k)ds.$$

The probability generating function Ψ for $z \in [0,1)$ as $\Psi(t,z) = \sum_{k\geq 0} z^k p(k,s,t)$, where the dependency on s is implicit.

Note that $z\frac{\partial\Psi}{\partial z} = \sum_{k\geq 0} kz^k p(k,s,t)$, $\frac{\partial\Psi}{\partial z} = \sum_{k\geq 0}(k+1)z^k p(k+1,s,t)$ and $z^2\frac{\partial\Psi}{\partial z} = \sum_{k\geq 0}(k-1)z^k p(k-1,s,t)$. Multiplying both sides of the equation in Proposition 3.3 by z^k and summing, we obtain after some simplifications, and writing $\nu = \lambda - \beta$:

$$\frac{\partial\Psi}{\partial t} = \left(\nu z^2 - (\nu+\mu)\,z + \mu\right)\frac{\partial\Psi}{\partial z}.$$

This first-order PDE has the boundary condition that at time $s \leq t$, there should be a single person in the cohort, so that $\Psi(s,z) = z$ since $p_n(1,1,s,s) = 1$. This equation is similar to the one solved in Baek, Kiet, and Kim (2007), although the variables are not the same: what they wrote as $\lambda(z-1)(z - \frac{\mu+\beta}{\lambda})$ corresponds here to $\nu(z-1)(z-\frac{\mu}{\nu})$, where we have defined ν as $\nu = \lambda - \beta$. Their solution for the equation is nevertheless valid and their approximation as well, so that

$$\lim_{t\to\infty} P(t,k) \sim k^{-\left(1+\frac{\lambda-\mu}{\nu-\mu}\right)},$$

which gives the result. ∎

This result implies that with a proportional growth, the name distribution converges towards a power law of parameter $1 + \frac{\beta}{\lambda-\beta-\mu}$. Further, if we write the limit $r = \lim_{t\to\infty}\frac{N(t)}{\rho(t)} = \frac{\beta}{\lambda-\mu}$, then the power law to which the distribution converges is fully conditioned by r and has paramater

$\frac{1}{1-r}$. In this model's framework, when there is no strategic name choice, the distribution of names is strictly related to the ratio of the number of names relative to the number of people in the population.

The results derived so far do not account for any strategy in naming. As we mentioned earlier, a term $\gamma(\rho, p, k-1)$ that could be introduced to represent the strategic choices would make the model intractable, if it effectively depends on these parameters. However, we can account for these strategic choices in a simpler manner. Indeed, based on Proposition 3.1, we argued that, depending on conditions such as the punishment cost c_w and the number of families I, strategic naming could be captured through either a choice of the rarest name, or population-weighed random choice of names. We will then define γ as the intensity of strategic choices: choices of the least common name if $\gamma > 0$, and choice of the most common names (population-weighed) if $\gamma < 0$. We then have the following:

Proposition 3.5 (Limiting Distribution with Strategic Naming)**.** *If $\gamma \neq 0$, the limiting distribution when t goes to infinity P_∞ verifies $P_\infty(k) \sim k^{2+\frac{\beta+\gamma}{\lambda-\beta-\gamma-\mu}}$. However, the ratio $\frac{N(t)}{\rho(t)}$ still converges to $\frac{\beta}{\lambda-\mu}$.*

Proof. The optimal strategy of chosing the least common name can in fact be simply represented: at any point in time, given the expected continuous increase in the population across all names, the least common name will be the latest invented name that has just been created. And if people strategically name their babies accordingly, the population with that name will increase, making the next new name the least common name. In a mean-field context, strategically choosing a name according to our model simply adds to the rate of name invention: by adding more people to the population carrying a name that was just invented, it makes the probability that this population increases by a given amount more likely.

If $\gamma < 0$ and represents the choice of the most common names, then in a mean-field context it is in fact not changing the distribution of the names with a positive population count, because the names that are redistributed are done so according to each name's share in the total population.

Hence, the parameter γ, whether positive or negative, which specifies the rate of strategic name choices, simply behaves like β: either it adds to it, reflecting the choice of brand new names, or it subtracts from it, reflecting the choice of existing names. The essential difference between the two is that γ does not add or take away new names, and therefore does not increase or decrease the name count: $\pi(s)$ does not depend on γ. In the proof for Proposition 3.5, the expression for $P(t, k)$ does not depend on β through the explicit form of $\pi(s)$, because it is cancelled out in the conditional probability $\mathbb{P}[A \in ds | A \leq t]$. As a result, replacing β with $\beta' = \beta + \gamma$ in Proposition 3.5 accounts for the impact of strategic name choices. Separately, β is not changed into β' in the expression of $pi(s)$, and the limiting behavior of $\frac{N(t)}{\rho(t)}$ is hence unchanged. □

Thanks to this result, we can now see that there is an additional degree of freedom in the relationship between the number of names per people, and the limiting behavior of the name distribution. Recalling that $r = \lim_{t \to \infty} \frac{N(t)}{\rho(t)}$, the name distribution therefore follows a power law of parameter

$$\frac{\lambda - \mu}{\lambda - \mu - \beta - \gamma} = \frac{1}{1 - r - \frac{\gamma}{\lambda - \mu}},$$

where γ may be positive or negative depending on characteristics of the strategic name choice.

4 Data Analysis

In this section we examine the empirical evidence concerning names in Ancient Greece, and confront it to the models we have discussed so far. We use the data from the LGPN[10], which was discussed in Gauthier (2021). However, since the LGPN covers more than a millenium with very significant political and social evolutions, it seems relevant to restrict it to the older archaic and classical periods, so we only retain observations anterior to 350 BC. Further, we know that the LGPN tracked names of both men and women, the latest representing only a fraction. In order to stay more strictly within the confines of our model, we only keep men names in our sample.

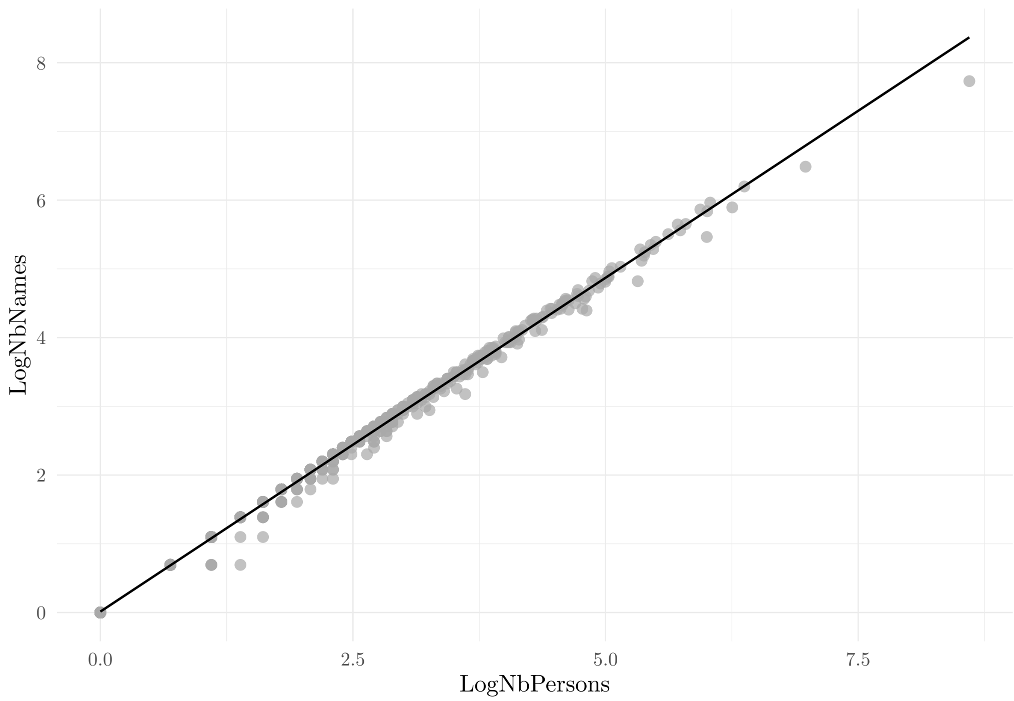
4.1 Number of Name Observations

As Baek, Kiet, and Kim (2007) reported, one typically observes a relationship between group size and the number of names observed in the group. The relationship is logarithmic in some cases (for contemporary China and Korea), or algebric with an exponent close to 1 (for many other countries), where using our notations $N = \rho^b$ with $b \in [0, 1]$.

Figure 5 shows on a logarithmic scale the relationship between the number of names and the number of people across *poleis*. We can see that it is quite strong. On the logarithmic scale, the slope in the data is 0.97; on the linear scale it is 0.46. Our simple model was predicting a simple linear relationship; the presence of this exponent slightly below 1 could be explained by a reduction in the attractiveness on large places, so that the immigration or name innovation rate tapers slightly off as population increases. In any case, the ratio on the linear scale tells us that

[10] See Parker, Yon, and Depauw (1996).

Figure 5: *Log/log Plot of the Number of Distinct Men Names as a Function of Total Number of Mentions Across Poleis*



the immigration / name innovation rate amounts to roughly half of the population growth.

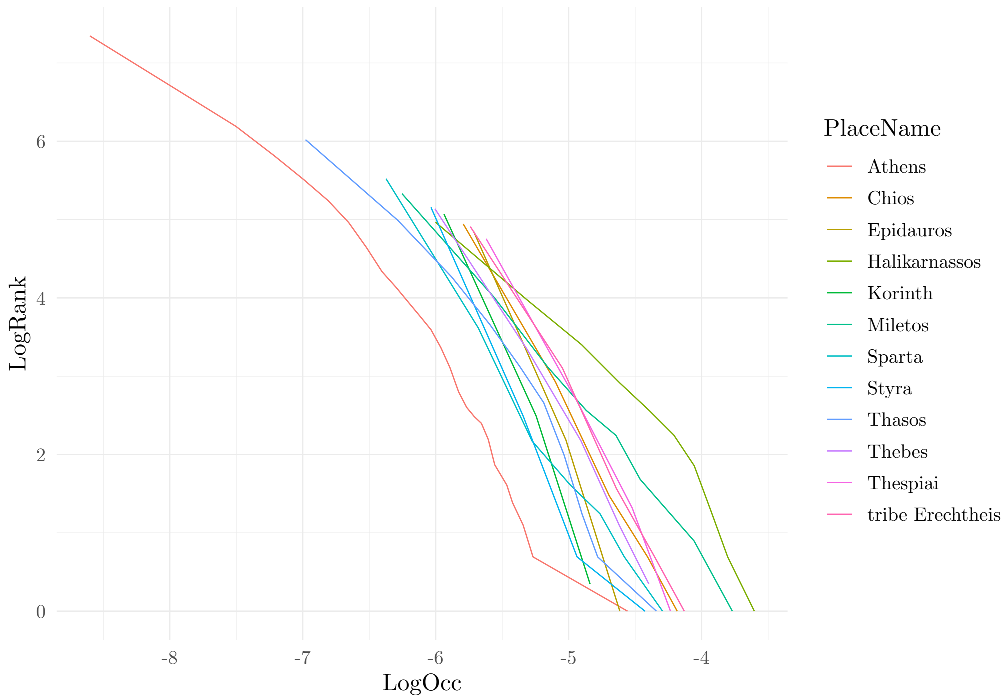
4.2 Empirical Name Distributions

Our formal model led us to expect a particular distribution shape, with particular parameters. Figure 6 shows the log/log plot of name frequency distributions across a large set of *poleis*. These distributions for the various places look quite similar, although they are somewhat shifted due to the differences in total count. We can see that they appear to follow power laws, but have a bend that could reflect an exponential distributino as well.

We can test for the various distributions that may explain these shapes. It is necessary to use a maximum-likelihood approach, and not a simple regression in log/log space, as was shown by Clauset, Shalizi, and Newman (2009). Table 1 shows the distribution fits for all the locations, carried out using the implementation from Alstott, Bullmore, and Plenz (2014). In aggregate, these distributions are power laws or truncated power laws. At the individual *polis* level, they look more like truncated power laws or exponential distributions, but the dispersion of parameters for the power law fits is much less dispersed than for these other types of distributions.

The power law coefficients, around 1.3, are quite different from what has been observed on contemporary data. As Baek, Kiet, and Kim (2007) reported, for most large cities or countries in the world, the coefficients for the name distribution power law tends to cluster around 2. Their

Figure 6: *Log/log Plot of Name Distributions Across Poleis for Men*



Note: The data includes all locations with more than 2000 observations.

Table 1: *Summary Statistics on Distribution Fits Across Places and in Aggregate*

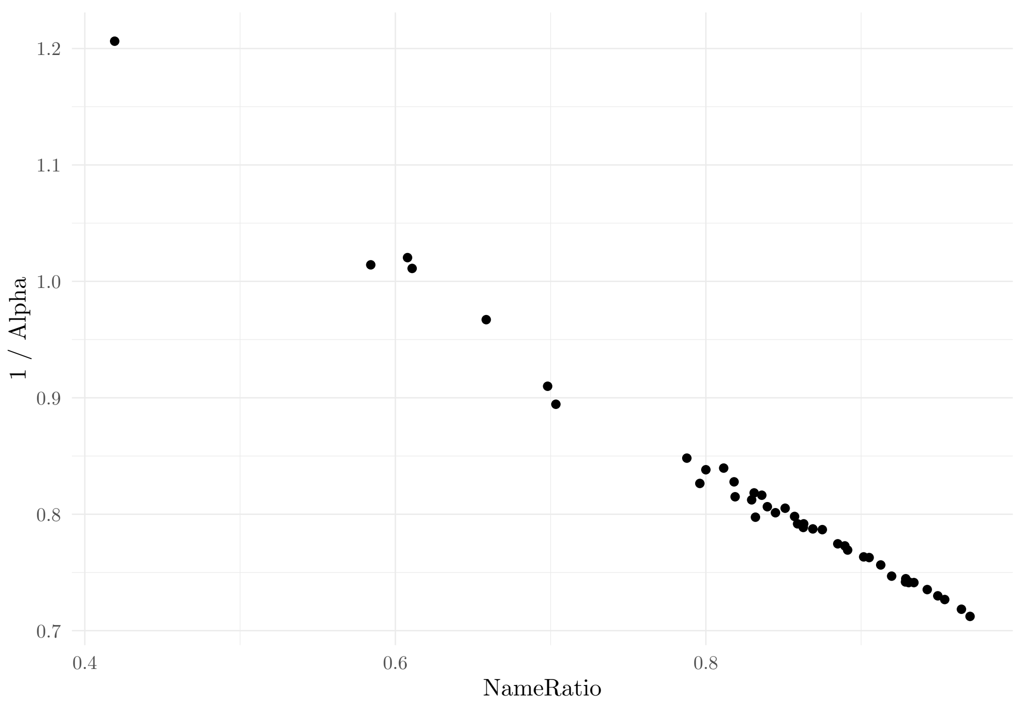
Statistic	Median	St Dev	Aggregate
Lambda Exp	1.959	0.669	0.444
Alpha Pow	1.263	0.124	0.832
Alpha Trunc	0.735	0.988	0.950
Lambda Trunc	0.801	0.725	0.017
Trunc vs Pow R	5.875	2.512	17.026
Trunc vs Pow p	0.000	0.000	0.000
Trunc vs Exp R	0.710	1.779	20.760
Trunc vs Exp p	0.141	0.274	0.000
Pow vs Exp R	-4.958	3.074	17.546
Pow vs Exp p	0.000	0.187	0.000

Note: The distribution names in the tests are abbreviated as follows: *Exp* = exponential, *Pow* = (pure) power, *Trunc* = power law with exponential decay. *R*: ratio of goodness-of-fit; a positive number means that the first law of the two is preferred. *p*: significance level; the probability that the preference would be due to randomness. The same abbreviations are used in other comparable tables.

model accounted for this pattern, and for the relationship with the name ratio $r = \frac{\beta}{\lambda - \mu}$. In a few particular cases, such as Korea, the power law has a parameter of 1, and the number of names grows logarithmically as a function of total population. One study, focused on rural Sicily, found coefficients that were closer to 1 and in some cases even under 1: Pavesi et al. (2003). Sicily's ancient name, *Magna Graeca*, reminds us that a share of its population came from Greece in the archaic period, so it is interesting that this particular shape in name distribution, where the parameter is close to 1, would be found there as well.

The models for the limiting behavior of name distributions developed by Baek, Kiet, and Kim (2007), Maruvka, Shnerb, and Kessler (2010) or Reed and Hughes (2003) effectively enforce a particular relationship between the name ratio and the power law parameter α, as we pointed out earlier: $\alpha = \frac{1}{1+r}$. Figure 7 plots the relationship in the case of Greece, across *poleis*, in logarithmic scale. We can see that it would not be possible to apply a relationship of the prescribed form. Adding the degree of freedom afforded by the strategic selection of names, we can express the relationship instead as $\alpha = \frac{1}{1+\frac{\gamma}{\lambda-\mu}-r}$ where in this expression $\gamma > 0$ shows the intensity of name choices of the conformist sort (picking the most common names).

Figure 7: *Relationship Between Distribution and Number of Names*



We find that $\frac{1}{\alpha} = 1.52 - 0.84\ r$, and the coefficient for r is fairly close to 1. In addition, we can see that $\frac{\gamma}{\lambda-\mu}$ = -0.52, which would suggest a fairly strong rate of conservative naming choices.

5 Conclusion

We began our analysis by asking why one would have names in the context of ancient Greece, and proposed a model inspired from identity and clan economics. We showed that the exchange of names could be understood as a way of improving social cooperation, in comparison with a full information case in particular. Given this context, we argued that optimal name choices could be reduced to some simple alternatives, for simplicity, and showed that adding the possibility for strategic names lead to particular distributions at the limit. Then, we examined empirical data from the archaic and classical periods, and showed that this model could well account for the patterns we observed.

References

Akerlof, George A., and Rachel E. Kranton. 2000. "Economics and Identity." *Quarterly Journal of Economics* 115 (3): 715–753.

Alstott, Jeff, Ed Bullmore, and Dietmar Plenz. 2014. "Powerlaw: A Python Package for Analysis of Heavy-Tailed Distributions." *PLoS ONE* 9 (1): e85777.

Baek, Seung Ki, Hoang Anh Tuan Kiet, and Beom Jun Kim. 2007. "Family Name Distributions: Master Equation Approach." *Physical Review E* 76 (4): 046113.

Barucca, Paolo, Jacopo Rocchi, Enzo Marinari, Giorgio Parisi, and Federico Ricci-Tersenghi. 2015. "Cross-Correlations of American Baby Names." *Proceedings of the National Academy of Sciences* 112 (26): 7943–7947.

Berg, Alastair. 2019. "Identity in Economics: A Review." SSRN Scholarly Paper ID 3421215. Rochester, NY: Social Science Research Network. https://doi.org/10.2139/ssrn.3421215.

Bernheim, B. Douglas. 1994. "A Theory of Conformity." *Journal of Political Economy* 102 (5): 841–877.

Bresson, Alain. 1981. "Règles de nomination dans la Rhodes antique." *Dialogues d'histoire ancienne* 7 (1): 345–362.

———. 2019. "Chapitre IV. Une famille camiréenne de commerçants en blé." In *La cité marchande*, 95–99. Scripta Antiqua. Pessac: Ausonius Éditions.

Busse, Thomas V., and Louisa Seraydarian. 1977. "Desirability of First Names, Ethnicity and Parental Education." *Psychological Reports* 40 (3): 739–742.

Carpenter, Jeffrey P. 2004. "When in Rome: Conformity and the Provision of Public Goods." *The Journal of Socio-Economics* 33 (4): 395–408.

Carr, Jack L., and Janet T. Landa. 1983. "The Economics of Symbols, Clan Names, and Religion." *The Journal of Legal Studies* 12 (1): 135–156.

Clauset, Aaron, Cosma Rohilla Shalizi, and M. E. J. Newman. 2009. "Power-Law Distributions in Empirical Data." *SIAM Review* 51 (4): 661–703.

Cline, Diane Harris. 2020. "Athens as a Small World." *Journal of Historical Network Research* 4 (May): 36–56.

Cuesta, Jose A., Carlos Gracia-Lázaro, Alfredo Ferrer, Yamir Moreno, and Angel

Sánchez. 2015. "Reputation Drives Cooperative Behaviour and Network Formation in Human Groups." *Scientific Reports* 5 (1): 7843.

Gaechter, Simon, Kyeongtae Lee, and Martin Sefton. 2020. "Risk, Temptation, and Efficiency in Prisoner's Dilemmas." 2020-15. *Discussion Papers.* Discussion Papers. School of Economics, University of Nottingham. https://ideas.repec.org/p/not/notcdx/2020-15.html.

Gauthier, Laurent. 2021. "A Distributional Perspective on Primary Sources from Ancient Greece." Working Paper hal-03315002v1. HAL. Paris: Université Paris 8. https://hal-univ-paris8.archives-ouvertes.fr/hal-03315002.

Gillen, Benjamin J. 2015. "A Note on Endogenous Norms in a Theory of Conformity." *Economics Letters* 129 (April): 57–61.

Hahn, Matthew W, and R Alexander Bentley. 2003. "Drift as a Mechanism for Cultural Change: An Example from Baby Names." *Proceedings of the Royal Society B: Biological Sciences* 270 (Suppl 1): S120–S123.

Hardie, Margaret M. 1923. "The Significance of Greek Personal Names." *Folklore* 34 (2): 149–154.

Herman, Gabriel. 1990. "Patterns of Name Diffusion Within the Greek World and Beyond." *The Classical Quarterly* 40 (2): 349–363.

Hornblower, Simon. 2010. *Thucydidean Themes.* Oxford: Oxford University Press.

Karila-Cohen, Karine. 2016. "Prosopographia Attica 2.0 : base de données et raisonnement prosopographique." *Revue historique* 680 (4): 869–904.

———. 2017. "Quand les familles s'affichent sur les pierres (à partir de l'exemple athénien)." *Pallas* Hors Série: 91–117.

———. 2018. "Le graphe, la trace et les fragments : l'apport des méthodes quantitatives et des outils numériques à l'étude des élites civiques athéniennes." *Annales HSS* 73 (4): 785–815.

———. 2019. "La valeur du lien : relations extérieures et capital social des notables athéniens de la basse époque hellénistique." In *La cité interconnectée dans le monde gréco-romain (IV^e^ siècle a.C.-IV^e^ siècle p.C.) : transferts et réseaux institutionnels, religieux et culturels aux époques hellénistique et impériale*, edited by Madalina Dana and Ivana Savalli-Lestrade, 287–306. Scripta antiqua 118. Bordeaux: Ausonius éditions.

Lee, Mi Jin, Woo Seong Jo, Il Gu Yi, Seung Ki Baek, and Beom Jun Kim. 2016.

"Evolution of Popularity in Given Names." *Physica A: Statistical Mechanics and Its Applications* 443 (February): 415–422.

Maruvka, Yosi E., Nadav M. Shnerb, and David A. Kessler. 2010. "Universal Features of Surname Distribution in a Subsample of a Growing Population." *Journal of Theoretical Biology* 262 (2): 245–256.

Mateos, Pablo, Paul A. Longley, and David O'Sullivan. 2011. "Ethnicity and Population Structure in Personal Naming Networks." *PLOS ONE* 6 (9): e22943.

Mengel, Friederike. 2018. "Risk and Temptation: A Meta-study on Prisoner's Dilemma Games." *The Economic Journal* 128 (616): 3182–3209.

Motschenbacher, Heiko. 2020. "Corpus Linguistic Onomastics: A Plea for a Corpus-Based Investigation of Names." *Names* 68 (2): 88–103.

Parker, Robert, Jean-Baptiste Yon, and Mark Depauw. 1996. "Lexikon of Greek Personal Names." 1996. http://www.lgpn.ox.ac.uk.

Parkin, Tim. 2013. "The Demography of Infancy and Early Childhood in the Ancient World." In *The Oxford Handbook of Childhood and Education in the Classical World*, edited by Judith Evans Grubbs and Tim Parkin, 40–61. Oxford University Press.

Pavesi, Angelo, Paola Pizzetti, Enzo Siri, Enzo Lucchetti, and Franco Conterio. 2003. "Brief Communication: Coexistence of Two Distinct Patterns in the Surname Structure of Sicily." *American Journal of Physical Anthropology* 120 (2): 195–199.

Puech, Vincent. 2012. "La méthode prosopographique et l'histoire des élites dans l'Antiquité tardive." *Revue historique* 661 (1): 155–168.

Reed, William J., and Barry D. Hughes. 2003. "On the Distribution of Family Names." *Physica A: Statistical Mechanics and Its Applications* 319 (March): 579–590.

Rogerson, Peter A. 2016. "The Effects of Months, Holidays, and the Birthdays of Presidents on Choice of Baby Names." *Names* 64 (4): 234–241.

Woods, Robert. 2007. "Ancient and Early Modern Mortality: Experience and Understanding." *The Economic History Review* 60 (2): 373.

Young, Robert K., Ana H. Kennedy, Ann Newhouse, Pam Browne, and Del Thiessen. 1993. "The Effects of Names on Perception of Intelligence, Popularity, and Competence." *Journal of Applied Social Psychology* 23 (21): 1770–1788.

Zanette, Damián H, and Susanna C Manrubia. 2001. "Vertical Transmission of Culture and the Distribution of Family Names." *Physica A: Statistical Mechanics and Its Applications*, Proceedings of the IUPAP International Conference on New Trends in the Fractal Aspects of Complex Systems, 295 (1): 1–8.