Business School

WORKING PAPER SERIES

Working Paper

2014-044

Hierarchical Organization and Performance Inequality: Evidence from Professional Cycling

Bertrand Candelon
Arnaud Dupuy

http://www.ipag.fr/fr/accueil/la-recherche/publications-WP.html

IPAG Business School
184, Boulevard Saint-Germain
75006 Paris
France

IPAG working papers are circulated for discussion and comments only. They have not been peer-reviewed and may not be reproduced without permission of the authors.

Hierarchical Organization and Performance Inequality: Evidence from Professional Cycling.*

Bertrand Candelon[†] and Arnaud Dupuy[‡]

November 4, 2011

Abstract

This paper proposes an equilibrium theory of the organization of work in an economy with an implicit market for productive time. In this market, agents buy or sell productive time. This implicit market gives rise to the formation of teams, organized in hierarchies with

*We thank Yin-Wong Cheung, Xavier Gabaix, Alfred Galichon, Michael Sattinger and Jesper van Thor and seminar participants at Ecole Polytechnique and Maastricht University for their valuable comments and Sander Dijksman for his data assistance.

†Department of economics, Maastricht University

‡Corresponding address: ROA and department of economics, Maastricht University PO Box 616, NL-6200 MD, The Netherlands. Email: a.dupuy@roa.unimaas.nl.

one leader (buyer) at the top and helpers (sellers) below. Relative to autarky, hierarchical organization leads to higher within and between team payoffs/productivity inequality. This prediction is tested empirically in the context of professional road cycling. We show that the observed rise in performance inequality in the peloton since the 1970s is merely due to a rise in help intensity within team and consistent with a change in the hierarchical organization of teams.

JEL Classification: D2, D3 and L22.

Keywords: Hierarchical organization, productive time, helping time, inequality, professional cycling.

1 Introduction

Many economies have witnessed a rising wage inequality in the last 5 decades (Acemoglu and Autor, 2010 and Acemoglu, 2003) alongside with large changes in many firms' organizational structure (see e.g. Rajan and Wulf, 2006). Early theoretical models developed by Lucas (1978) and Rosen (1982) exhibit that earnings inequality raises with span of control. This prediction has recently been tested empirically in a few studies. Fox (2009) shows that earnings inequality increases with job responsibility in Swedish and US firms, and Gabaix and Landier (2008), Garicano and Hubbard (2009) and Tervio (2008) conclude that the recent increase in earnings inequality in large US firms and in law firms is largely due to the rise in span of control in these firms.

This paper proposes an alternative channel through which hierarchical organization and earnings inequality might be related, namely help intensity. This paper studies an economy with an implicit market for productive time. The scarce resource in this market is the time each agent can dedicate to production. Agents can dedicate their productive time to their own production (autarky), buy (part of the) productive time from helpers and herewith increase their own production, or sell (part of) their productive time to a leader, hence giving up own production. Since helping time of better helpers is more efficient, the hedonic equilibrium price for productive time compensates i) forgone own production and ii) helping ability. This implicit market for productive time gives rise to teams' formation. These teams have a hierarchical organization with a leader at the top producing output with the help of helpers below.

There are many examples of goods and services produced with teams organized that way. We may think for instance of a lawsuit: The defence of a case is performed by a leading lawyer who receives full credit for the outcome of the trial. The leading lawyer however might receive help from other lawyers at her firm to prepare the trial. The opportunity cost of spending productive time helping the leading lawyer requires a compensating wage. Yet, other examples can be found in architect offices, maisons de Haute Couture, music bands and many more... Interestingly enough, certain sports, in particular professional road cycling, show a similar organization of teams but present the advantage of having the size of teams fixed exogenously. These economies allow us to abstract from the extensive margin and focus on the core: the relationship between helping time and performance inequality.

The existence of an implicit market for productive time raises many interesting questions about the structure of such an economy. For instance, who becomes a leader and who becomes a helper? Are leaders more able than helpers? Are more able helpers assigned to more able leaders? How does the distribution of payoffs look like? Is there more inequality in payoffs with hierarchical organization than with autarky?

In this paper, we elaborate a core theoretical model using stylized features of professional road cycling. In this core model, the distribution of roles (leaders, helpers and self-production), the assignment of leaders to helpers, and the distribution of payoffs (the hedonic price for productive time) are endogenously determined. We show that under mild conditions, an equilibrium in this economy exists and is Pareto optimal. The equilibrium assignment is so that i) within teams the better rider becomes the leader, ii) better helpers are matched with better leaders and iii) some form of stratification arises: more able agents either become leaders or helpers and less able agents either become helpers or ride individually. The model can be used to study the link between hierarchical organization and productivity and/or payoffs inequality. For instance, we show that relative to autarky, hierarchical organization leads to greater payoffs/performance inequality.

The core model developed in this paper is most closely related to the one-sided assignment[1] models studied by Lucas (1978), Rosen (1982), Garicano and Rossi-Hansberg (2004;2006) and Garicano and Hubbard (2005). However, in contrast with our model, in these models, output is produced by agents at the bottom of the hierarchy helped by managers at the top of the hierarchy. In the economy studied by Garicano and Rossi-Hansberg (2006) for instance, hierarchical organization arises as an efficient way of sharing knowledge. In contrast with our model, the implicit market is therein a market for knowledge, not productive time. Agents can buy knowledge either directly by learning, which is costly, or indirectly by "hiring" more knowledgeable agents to solve problems they cannot solve themselves. In these hierarchical organizations, output is produced by workers helped by more knowledgeable agents called managers.

In Garicano and Rossi-Hansberg (2006), as in Lucas (1978) and Rosen (1982),[2] a strict stratification arises with production workers at the bottom of

[1] This contrasts with two-sided assignment models studied in Tinbergen (1956), Becker (1973), Rosen (1974) and Sattinger (1993) among others where agents on one side of the market (workers, women, sellers) meet agents on the other side of the market (firms, men, buyers). These models have been used in the recent literature on CEO pay by e.g. Gabaix and Landier (2008) and Tervio (2008). More recently, Edmans, Gabaix and Landier (2009) embedded a moral hazard problem into a talent two-sided assignment model to study CEO-pay incentives.

[2] Rosen (1982) recognizes that if the rents function is sufficiently convex then the model predicts that both the least and most able agents become managers. Rosen (1982) classifies this case as a "pathology" of the model since it implies that some managers are less able than their workers.

the distribution of ability, self-employed in the middle and managers (helpers) at the top. Our model allows for a more general form of stratification in equilibrium: while leaders are more able than self-employed, some helpers may be more able than some leaders (of different helpers). Similarly, self-employed agents could be less or more able than helpers depending on the primitives of the model.

This paper also relates to the literature on the Monge-Kantorovich transportation problem (see e.g. Villani, 2009). Recently, Chiappori, McCann and Nesheim (2010) have shown equivalence results between two-sided quasi-linear hedonic models and the Monge-Kantorovich optimal transportation problem. We use existence and duality results from the Monge-Kantorovich transportation literature (Villani, 2009) to study the properties of the one-sided hedonic economy presented in this paper. In particular, we build on Chiappori, Galichon and Salanie (2011) that show how a one-sided assignment model can be re-formulated as a two-sided assignment model with symmetric surplus, to prove that a competitive equilibrium in our model is Pareto optimal.

Anticipating on our empirical results, we show that, since the 1970s, the help intensity within teams has increased sharply in the peloton of the Tour de France, an increase that is related to changes in the incentives to organize hierarchically within teams. In turns, we show that the observed rise in performance inequality in the peloton since the 1970s is merely due to this increase in help intensity.

The remainder of the paper is organized as follows. Section 2 proposes a theoretical model that introduces hierarchical organization arising from the existence of an implicit market for productive time. Section 3 studies the example of professional road cycling and in particular the Tour de France. Section 4 concludes.

2 Theoretical model

2.1 Set up

We present the model using the analogy to the Tour de France but one could also consider any of the previous examples (lawsuit, haute couture, rock band,...). In these cases, one can replace "riders" by "agents" throughout this section.

Let riders be endowed with one unit of time and with ability z, $z \in Z$

where $Z = [\underline{z}; \overline{z}]$ with $0 < \underline{z} \leq \overline{z} < \infty$.[3] Let μ be the probability measure representing the distribution of riders' ability on Z. We assume that ability is measured in terms of velocity (kilometers per hour) that is, z is the total distance of the Tour divided by the time it would take rider z to cover this distance when riding individually.

In this economy, riders decide either to ride individually or coordinate their efforts within teams. The number of riders in a team is fixed exogenously by rules and for the sake of simplicity we assume it is equal to 2.[4] Within each team, one rider becomes the helper of the other. The two riders must decide who becomes the leader and who becomes the helper. The distribution of roles within teams is endogenous to the model. Conditions under which the most able rider becomes the leader are given below in Proposition 3.

In case riders decide to coordinate their efforts, the helper will devote $s \in [0, 1]$ unit of time to her leader. Let $v_h(z_h, s) = z_h - saz_h$ be the velocity of a helper of ability z_h when providing helping time s to her leader. The

[3] We could work with unbounded support for z without changing much of the model but this would require to write down boundary conditions for the functions f and v in our standing assumptions below.

[4] This assumption restricts the extensive margin compared to Garicano and Rossi-Hansberg (2006) who study leverage in a knowledge economy. However, this does not seem to be restrictive in our case since the number of riders in each team participating in the Tour is fixed to 9 by rules.

parameter a defines a lower bound for the helpers velocity: a full-time helper has velocity $z_h(1-a)$.[5] The term az_h indicates the helper's velocity loss by unit of helping time provided to her leader.[6] It seems natural that this loss will depend on the ability of the helper since, by symmetry, az_h can also be interpreted as the helper's velocity gain by unit of time spent helping himself. The helper's forgone velocity requires a compensation from her leader, say $w_h(z_h, s)$. This compensation depends on s and z_h since the forgone velocity depends on helping time provided s and helper's ability z_h.

Similarly, let the velocity of a leader with ability z_l helped by a helper of ability z_h providing helping time s be $v_l(z_l, s) = z_l + f(z_h)s$ with $(z_l, z_h) \in Z^2$, where $f(z_h)$ is a twice differentiable continuous function. The function $f(z_h)$ indicates a leader's velocity gain by unit of helping time provided by a helper of ability z_h.[7] This velocity gain generates higher outcome that are (partly)

[5]For $a = 1$, $v_h(z_h, 1) = 0$. This is unrealistic since during the Tour, riders that do not finish within a certain interval of time after the stage winner are disqualified, i.e. $v_i(z_i, s) > v_{\min}$ where $v_{\min}$ is the velocity below which riders are disqualified. This rule is either a constraint on a or a constraint on the helping time riders can provide. In the remaining of the paper, we assume for simplicity that a is low enough, $a < 1$, so that $v_h(z_h, 1) > v_{\min}$.

[6]Note that all that matters in this economy is the relative shape of the loss of helpers' velocity to the gain in leaders' velocity. Assuming a linear shape for the loss in helpers' velocity, i.e. az, is without loss of generality since the properties of the gain in leaders' velocity, i.e. the function $f()$ below are defined relative to a.

[7]Note that the properties of the economy presented in the propositions below continue to hold if we allow f to depend on the leader's ability as long as the following restrictions are satisfied: i) $\frac{\partial f(x,y)}{\partial x} \geq 0$, ii) $\frac{\partial^2 f(x,y)}{\partial x \partial y} \geq 0$ and iii) $f(x,y) - ay \geq f(y,x) - ax$ for all

used to compensate the helper. Note that if $s = 0$, even though riders are in the same team, both riders bike individually and their respective velocity is simply $v_l(z_l, 0) = z_l$ and $v_h(z_h, 0) = z_h$.

Throughout this paper we maintain the following assumptions about the efficiency of helping time.

Condition 1 *Standing Assumptions I (SA I, hereafter)*

1. Helping time is strictly efficient for the leader's velocity, $f(y) > 0$ for all $y \in Z$ but strictly costly for the helper's velocity, $1 > a > 0$,

2. Better riders are also better helpers, $f'(y) > 0$ for all $y \in Z$.

3. Inefficiency of helping someone else increases with ability, $f(y) - ay \geq f(x) - ax$, for all $(x, y) \in Z^2$ with $x > y$.[8]

Assumption SA I.3 has a strong intuitive interpretation. Remember that az can be interpreted as rider z's velocity gain by unit of time spent helping

$x \geq y$. In words: i) requires that better leaders make better use of their helper's time and ability, ii) the ability of leaders and helpers are complementary and iii) the inefficiency of helping someone else than oneself increases with ability. For the sake of expositional simplicity, we present in this paper the special case where $f(x, y) = f(y)$. Proofs of the propositions in the general case are available upon request from the authors.

[8]Note that writing $x = y + h$, after some rewritting, the constraint becomes $\frac{f(y+h)-f(y)}{h} \leq a$. This must hold for all $h > 0$ and all $y \in Z$. Hence, by definition, we have $\lim_{h\to 0} \frac{f(y+h)-f(y)}{h} = f'(y) \leq a$ for all $y \in Z$.

herself. Similarly, $f(z)$ is the velocity gain by unit of time generated by a rider of ability z when helping someone else. The difference $f(z) - az$ is therefore interpreted as the inefficiency, in velocity units, of helping someone else than oneself. Assumption SA I.3 requires that this inefficiency increases with the ability of riders. Stated otherwise, Assumption SA I.3 implies that it becomes increasingly difficult to communicate the subtleties of the activity with someone else.[9]

We assume that the demand for performance is exogenous to the model so that velocity is priced exogenously in the market. Let $p(v)$ be a continuous and twice differentiable function mapping rider's velocity v into money prizes.

Throughout this paper we maintain the following assumptions about the reward function.

Condition 2 *Standing Assumptions II (SA II, hereafter))*

1. $p(0) = 0$,

[9] To our knowledge, there is no evidence either theoretical or empirical about the shape of this efficiency. However, a link exists with the psychology literature on the effect of multitasking on productivity. −One may indeed think of helping someone else as multitasking since it requires, in addition to the action of helping (thinking about how to increase productivity), interacting and communicating with someone else.− In particular, Rubinstein, Meyer and Evans (2001) show that agents lose time when they have to switch from one task to another and that these "time costs" increase with the complexity of the task. We thank Patricia Crifo for pointing out this literature.

2. Rewards $p(.)$ are strictly increasing in velocity $p'(z) > 0$,

3. Rewards $p(.)$ are convex in velocity, i.e. $p''(z) \geq 0$ for all $z \in Z$,

4. $0 < p'(z) < \infty$ for all $0 < z < \infty$.

Note that the assumption of convexity of the reward function is more than supported by empirical data on the distribution of prizes by rank in the final classification, see Figure 1.

Let Y, the surplus of a team, be given by the sum of all prizes won by its riders. Formally, let $Y(z_l, z_h, s) = p(z_l + f(z_h)s) + p(z_h(1 - as))$.[10] Let $w_h(z_h, s)$ be the payoffs of a helper z_h providing helping time s and let $w_l(z_l, s)$ be the payoffs of a leader z_l enjoying s helping time from her helper. Without managers, total surplus is split among the riders so that $Y(z_l, z_h, s) = w_h(z_h, s) + w_l(z_l, s)$.

2.2 Feasible teams

Some general results about feasible teams are helpful to characterize the equilibrium of this model. First, it can be shown that Proposition 3 is true under SA I and SA II.

[10] It is interesting to note that the ability of riders are complementary in surplus if and only if they help each other, i.e. $\frac{\partial^2 Y}{\partial z_h \partial z_l} = p'' f' s > 0$ if and only if helping time $s > 0$.

Proposition 3 *Under SA I and SA II, the surplus of all feasible teams* $(z_l, z_h) \in Z^2$*, with* $z_l \geq z_h$*, is maximized when* z_l*, i.e. the most able rider, becomes the leader and* z_h*, i.e. the least able rider, becomes the helper.*

Proof. See Appendix B ■

The intuition is the following. From SA I.3, $ay - f(y) \leq ax - f(x)$ for $x > y$. This means that the net gain of velocity in a team $\langle y, x\rangle$ is larger if the best rider is re-assigned from being the helper to being the leader. Hence, as long as rewards are convex as stated in SA II.3, a team's surplus is greater when the less able rider helps the most able one. Denoting $\gamma(z_l, z_h)$ the measure connecting helpers to leaders in equilibrium, from Proposition 3 we already know that $d\gamma(z_l, z_h) = 0$ for $z_l < z_h$.

Another interesting pattern of the model is the strategy within teams. How much help intensity to ask/offer? To answer this question, first note that, within teams, riders will always choose s so as to maximize their team's surplus, i.e. $s^*(z_l, z_h) = \arg\max_s Y(z_l, z_h, s)$.[11] The arguments of s^* will be dropped when unambiguous.

[11] Indeed, suppose that both riders choose $s^0 \neq s^*(z_l, z_h)$ and that the helper receives payoffs w_h^0 and the leader payoffs $w_l^0 = Y(z_l, z_h, s^0) - w_h^0$. Since by definition $Y(z_l, z_h, s^0) \leq Y(z_l, z_h, s^*(z_l, z_h))$, both riders could increase their team's surplus by setting $s = s^*(z_l, z_h)$. Splitting the additional surplus $Y(z_l, z_h, s^*(z_l, z_h)) - Y(z_l, z_h, s^0)$ among them will increase both riders' payoffs. For all feasible teams, we therefore always have $s = s^*(z_l, z_h)$.

Under our standing assumptions, the following proposition shows that the decision about how much helping time to provide simplifies considerably for all feasible teams.

Proposition 4 *Under SA I and SA II, for all feasible teams* $(z_l, z_h) \in Z^2$*, with* $z_l \geq z_h$*, surplus* $Y(z_l, z_h, s)$ *is strictly convex in* s *such that* $s^*(z_l, z_h) = \left\{ {}^{1 \text{ iff } Y(z_l,z_h,1)>Y(z_l,z_h,0)}_{0 \text{ otherwise}} \right.$.

Proof. See Appendix B ■

Intuitively, the convexity of the reward function in assumption SA II.3 carries on to the relationship between teams' surplus and helping time as long as helping time is efficient and $a > 0$ as in SA I.1. In other words, if p is sufficiently convex given f and a, then for all feasible teams $s^*(z_l, z_h) = 1$. Reciprocally, if f is sufficiently small relative to a given p, then for all feasible teams $s^*(z_l, z_h) = 0$.

Finally, for notational convenience, since from Proposition 4, $s^* = 1$ if $Y(z_l, z_h, 1) > Y(z_l, z_h, 0)$ and 0 otherwise, we define $Y_1(z_l, z_h) \equiv Y(z_l, z_h, 1) = p(z_l + f(z_h)) + p(z_h(1 - a))$ and $Y_0(z_l, z_h) \equiv Y(z_l, z_h, 0) = p(z_l) + p(z_h)$. Similarly, let $w_i(z) = w_i(z, 1)$ for $i = h, l$, and $w_0(z) = w_i(z, 0)$ for $i = h, l$ be the payoffs of individual riders. Note that $w_0(z) = p(z)$ independently of the ability of her team mate.

2.3 Riders' problem

Riders maximize their payoffs. The problem of a rider z is therefore to choose the role (leader, helper or individual rider) that maximizes her payoffs:

$$\max\left\{w_l(z), w_h(z), p(z)\right\}.$$

The leader's problem is to find a helper z_h that maximizes her payoffs. This problem reads as: $\max_{z_h}\left[Y_1(z_l, z_h) - w_h(z_h)\right]$. The first order condition to the leader's problem yields:[12]

$$w_h'(z_h) = p'(z_l + f(z_h))f'(z_h) + (1-a)p'(z_h(1-a)) > 0. \qquad (1)$$

From our standing assumptions SA I and SA II we already know that the equilibrium payoffs function for helpers is strictly increasing in helpers' ability.[13]

Symmetrically, the helper's problem is to find a leader z_l that maximizes her payoffs. This problem reads as $\max_{z_l}\left[Y_1(z_l, z_h) - w_l(z_l)\right]$. The first order

[12] We refer the reader to Appendix A for a formal presentation of the second order conditions to the leader's and helper's optimization problem.

[13] As shown in Appendix A, the second order conditions to the riders' problem also indicates that the payoffs functions are strictly convex.

condition to this problem is:

$$w_l'(z_l) = p'(z_l + f(z_h)) > 0. \tag{2}$$

From our standing assumptions SA I and SA II, it follows that the equilibrium payoffs function for leaders is strictly increasing in leaders' ability.

The first order conditions pin down the slopes of the payoff functions. The level of these two functions $w_l(\underline{z})$ and $w_h(\underline{z})$ together with the slopes will determine the set of riders for which $w_h(z) = \max\{w_l(z), w_h(z), p(z)\}$ (the set of helpers), the set of riders for which $w_l(z) = \max\{w_l(z), w_h(z), p(z)\}$ (the set of leaders) and the set of riders for which $p(z) = \max\{w_l(z), w_h(z), p(z)\}$ (the set of individual riders).

2.4 Equilibrium

Let $L \subseteq Z$ be the set of leaders, $H \subseteq Z$ the set of helpers and $I = Z \backslash L \cup H \subseteq Z$ the set of individual riders. Note that the sets H and L are endogenously determined and not necessarily disjoint. This allows for the possibility that at some ability level, riders might be indifferent between being a leader or a helper.

Definition 5 *A feasible assignment in this economy is a positive measure γ on Z^2 such that:*

$$\int_{z_l \in Z} d\gamma(z_l, z) + \int_{z_h \in Z} d\gamma(z, z_h) = d\mu(z), \forall z \in Z.$$

Intuitively, $\int_{z_l \in Z} d\gamma(z_l, z)$ is the quantity of riders z that become helpers while $\int_{z_h \in Z} d\gamma(z, z_h)$ is the quantity of riders z that become leaders. Let $\Gamma(\mu)$ be the set of feasible measures γ given μ.

Definition 6 *An equilibrium in this economy consists of:*

- *two payoffs functions $w_h(z_h)$ and $w_l(z_l)$ and,*
- *a feasible assignment γ and sets $L \subseteq Z$ and $H \subseteq Z$ such that: riders choose a role (helper, leader or individual rider) and their eventual leader or helper so as to maximize their own payoffs.*

The following two propositions show important characteristics of the equilibrium assignment.

Proposition 7 *Under SA I and SA II, in equilibrium, more able leaders are matched with more able helpers.*

Proof. See Appendix B ■

This result essentially follows from the convexity of the reward function p and the fact that better helpers also happen to be better riders.

Proposition 8 *Let $z^{(l)}$ be the ability of riders such that $w_l(z^{(l)}) = p(z^{(l)})$. Under SA I and SA II, there is the following stratification in equilibrium: i) there are no leaders of ability lower than $z^{(l)}$ and ii) there are no individual riders of ability higher than $z^{(l)}$.*

Proof. See Appendix B ■

This type of stratification is much more general than the strict stratification obtained in other models of organization. For instance, the model developed by Rosen (1982) predicts that more able agents become managers of less able ones (workers). Garicano and Rossi-Hansberg's (2006) model is a bit richer in terms of stratification as it allows for self-employed workers. However, still the least able manager is more able than the most able self-employed and the least able self-employed is more able than the most able worker. The model presented in this paper offers a richer stratification where, interestingly enough, some leaders could be of lower ability than some helpers (of more able leaders), some individual riders could be of lower ability than some helpers and at some ability level, some riders may become leaders, others helpers.

Several special cases are worth noting. First, suppose that $z^{(l)} > \overline{z}$. In that case, all riders prefer riding individually than becoming a leader. Since, there are no leaders there cannot be any helper either. All riders ride individually. This case arises when $p(.)$ is not convex enough given $f(.)$ and a.

Second, suppose that $z^{(l)} < \underline{z}$. This means that all riders prefer becoming a leader than riding individually. There are no individual riders.

Finally, as soon as $z^{(l)} < \overline{z}$, there will be intervals of helpers and leaders. In particular, a perfectly stratified equilibrium could arise. We could have for instance riders of ability $z \in \left[\underline{z}, z^{(h)}\right)$ becoming individual riders, riders $z \in [z^{(h)}, z^{(l)})$ becoming helpers and riders of ability $z \in [z^{(l)}, \overline{z}]$ becoming leaders. Similarly, we could have a stratification where riders of ability $z \in [\underline{z}, z^{(h)})$ become helpers, riders of ability $z \in [z^{(h)}, z^{(l)})$ become individual riders and riders of ability $z \in \left[z^{(l)}, \overline{z}\right]$ become leaders.

2.5 Existence and optimality

To study the properties of an equilibrium in this economy it is useful to write the social planner's problem (SPP) associated. Using that $Y_0(z_l, z_h) = p(z_l) + p(z_h)$ and the constraint in Definition 5, this problem reads as:

$$SPP^* = \max_{(L,H)\in Z^2, \gamma\in\Gamma(\mu)} \left\{ \int_I p(z)d\mu(z) + \int_{L\times H} Y_1(z_l, z_h)d\gamma(z_l, z_h) \right\} \quad \text{(P1)}$$

We first show that an optimal solution (L, H, γ) exists.

Proposition 9 *Under SA I and SA II, an optimal solution* (L, H, γ) *to* SPP *exists.*

Proof. See Appendix B ■

Defining $w := \max\{w_l, w_h\}$ for notational convenience, the dual program (D1) to the social planner's problem is:

$$\begin{aligned} DP^* &= \min_w \int_Z w(z)d\mu(z) && \text{(D1)} \\ & s.t. \\ w(z) &\geq p(z) \text{ for all } z \in Z \text{(i)} \\ w(z_l) + w(z_h) &\geq Y_1(z_l, z_h) \text{ for all } z_l, z_h \in Z^2 \text{ (ii)} \end{aligned}$$

Constraint i) in the dual program corresponds to the condition of individual rationality (riders always have the option of remaining unmatched). Constraint ii) guarantees that an outcome is not blocked by any coalition. If ii) is not satisfied, a pair z_l and z_h can always break up and reunite splitting $Y_1(z_l, z_h)$ in such a way that both are better off (z_l gets more than $w(z_l)$ and z_h gets more than $w(z_h)$). Using constraint i) and ii), it is easy to see that the weak duality inequality holds $DP^* \geq SPP^*$. Proposition 10 in fact states that there is duality.

Proposition 10 *Under SA I and SA II, there is duality, i.e.* $SPP^* = DP^*$.

Proof. See Appendix B ■

This result allows us to prove the following proposition.

Proposition 11 *A feasible tuple* $((w_h, w_l), (L, H, \gamma))$ *that solves both the primal and dual program maximizes riders payoffs.*

Proof. See Appendix B ■

A direct corollary of Proposition 11 is:

Corollary 12 *The pair of payoffs functions* (w_l, w_h) *that maximizes riders' payoffs is Pareto Optimal.*

Proof. See Appendix B ■

2.6 Comparative statics

Performance inequality will be lowest in autarky, measuring inequality by the range. This result is obvious for two reasons. First, from Proposition 3, in a hierarchically organized team, the most able rider becomes leader of the team. Second, from Proposition 8, riders at the top of the ability distribution become either leaders or helpers. Combining these two results, in an economy with hierarchical organization, the most able riders must necessarily be leaders. The performance at the top of the distribution in the economy with hierarchical organization must therefore be larger ceteris paribus than in autraky. Similarly, from proposition 8, we know that riders at the bottom of the distribution of ability either become a helper or ride individually. Hence ,the performance at the bottom of the distribution of ability is at most equal to that of a ceteris paribus same economy in autarky. It follows that the range will be lowest in autarky ceteris paribus.

Without imposing further structure, the model also makes predictions about payoffs inequality. Since the best riders become leaders and since $w_l^{''} > p^{''} > 0$ we know that payoffs inequality in the upper tail is larger in the economy with hierarchical organization. We also know that the least able riders either become individual riders or helpers. However, since, without more structure, we cannot conclude about the sign of $w_h^{''} - p^{''}$, we cannot conclude about inequality in the lower tail.

Regarding the range of payoffs, if the least able rider is an individual rider, we have $w_l(\overline{z}) \geq p(\overline{z})$ and hence $w_l(\overline{z}) - p(\underline{z}) \geq p(\overline{z}) - p(\underline{z})$ such that the range must be lower in autarky. However, if the last rider is a helper, we have $w_h(\underline{z}) \geq p(\underline{z})$ and $w_l(\overline{z}) \geq p(\overline{z})$, and we cannot conclude about $w_l(\overline{z}) - w_h(\underline{z}) \lesseqgtr p(\overline{z}) - p(\underline{z})$.

3 Empirical evidence

3.1 Organization in road cycling

Testing the predictions of our model requires at the minimum to have access to data about an industry where helper-leader relationships exist (condition 1). Measuring the impact of hierarchical organization on performance inequality in the industry requires in addition that i) the performance $v(z)$ is measured at the individual level (condition 2), ii) the industry has experienced a change in the incentives to organize hierarchically within teams (condition 3) and, iii) a measure of the extent of hierarchical organization within teams is available (condition 4). Road cycling constitutes a rare case where these 4 requirements are satisfied.

Condition 1: helper-leader relationship.

Road cycling racing *is* an individual sport where the first to cross the finish line wins. However, unlike most individual sports, road cycling riders have traditionally been grouped into teams. Teams' tactics (organization) has become an important aspect of the sport. Tactics turns out to be inherent in this sport since the aerodynamic benefit of drafting, following as closely as possible the slipstream of the rider in front, can save as much as 40% of the energy compared to riding alone. Some teams therefore designate a leader and have the remaining riders serve as a "wind shield" for their leader to spare energy until critical moments of the race (final climb during a mountain stage for instance). Helpers also play the role of a "donkey" during races, carrying food and water to their leader, or exchange their wheels or even bike in case of a mechanical problem of their leader during the course. This hierarchical organization is especially important during stage races, among which The Tour de France constitutes the event of the year.[14]

[14] See McGann and McGann (2006) for a detailed account of the history of the Tour de France.

In 2008, the winner of the Tour de France received €450,000 in prize money or 5 times more than the winner of the Giro (€90,000). That year, the total prize money distributed on the Tour de France amounted to €3,269,760 or 2.4 times more than on the Giro (€1,380,010).

Condition 2: measure of individual performance.

The performance of each rider $v(z)$ is readily measured by his average velocity defined as the total distance of the event divided by his finishing time.

Condition 3: change in organizational incentives.

It is worth noting that since the end of the 1960s, several developments indicate that teams participating at the Tour de France have progressively become more hierachically organized, respecting hence condition 3.

First, the Tour de France has moved from national to trade teams at the end of the 1960s. Initially, the Tour was opened to all riders and most of them were enrolled in trade teams. However, the organizer of the Tour, Henri Desgrange, insisted that while riders could compete in the name of their sponsors, no cooperation or tactics would be allowed between these riders. However, in 1929, the Belgian rider Maurice De Waele won the Tour with the "illegal" help of his team mates even though he was ill. This event marked Henri Desgrange: "My Tour has been won by a corpse," and led him to deny participation to trade teams. Only national and regional teams were allowed from 1930 until 1961. The hierarchical organization within teams became more difficult since most riders belonged to rival trade teams for the rest of the season. The loyalty of riders was sometimes questionable, within and between teams, leading to an inefficient organization as can testimony several famous events.[15] Under the pressure of sponsors that paid the salaries of riders the whole year long but were denied publicity from the season's major event,[16] the organizers decided to come back to trade teams by the end of the 1960s for good.

Second, the media exposure of the Tour de France has grown ever since and so have the stakes. Figure 2 shows the evolution of the total prize money, corrected for inflation, distributed on the Tour since 1950. The total prize money were roughly steady from 1950 to 1971 and started increasing ever since, at 3% per year between 1971 and 1985 and 5% thereafter.

Third, we observe a convexification of the payoffs by rank. As shown in Figure 3, the average share of prizes allocated to the winner was about 4.5% between 1950 and 1975. It increased to an average of 6.8% between 1975 and 1985. Since 1985, the winner goes home with about 15% of the total amount

[15] In 1959, the French team was made up of many strong riders such as Raphaël Géminiani, Henri Anglade and Jacques Anquetil. The French team was full of internal rivalries. Part of team decided to help spanish rider Federico Bahamontes win rather than Henri Anglade in the hope to win more fees during the post-Tour criteriums as Bahamontes was a much poorer rider on flat closed circuits than Anglade.

[16] Trades were partly accommodated for with the authorization for the riders to put their respective trades name on their jersey and the introduction of the Caravan. The Caravan consists in a trade parade preceding the riders during the Tour de France. This caravan is praised by the spectators and reached an height between the 30s and the 60s.

of money prizes distributed during the Tour de France.[17] A prediction of our model is that a convexification of the payoffs function $p(.)$ leads to greater incentives to organize teams hierarchically.

Condition 4: measures of the extent of hierarchical organization within teams.

A key element of the Tour de France to measure the extent of hierarchical organization, is that a few of the 20 stages that determine the final classification are individual time trials. An individual time trial is a stage during which riders ride alone against the clock. There is no help possible between riders during such a stage such that $s = 0$ for all teams during these stages. This means that the velocity of a rider during such a stage reflects his true individual ability.[18] It follows that the difference in velocity between the leader of a team and his helpers during an individual time trial reflects a team composition effect, i.e. the pure ability differences between riders. This

[17]It should be noted that the prizes won by each rider of a team are usually pooled together and redistributed within the team.

[18]It might be argued that helpers will put on less effort during a long individual time trial to save energy for helping their leader in the next stages. Fortunately, since 1970, most Tours started with a short individual time trial, i.e. the Prologue. Compared to other individual time trials that generally last between 45 to 60 minutes and occur after several stages, the Prologue is a short effort of about 10 to 15 minutes that occurs before any other stages. This means that riders are fresh from the start and recover rapidly from their efforts during the Prologue. Furthermore, the stages following the Prologue are generally flat stages ending with a massive sprint so that the amount of helping time devoted by helpers of final classification riders is limited during the stages following the Prologue. Hence, in contrast to other individual time trials, during the Prologue, riders have no incentives not to perform at their best, and for each rider the measured velocity during the Prologue reveals the true ability of riders $v(z) = z$.

contrasts with the velocity difference in the final classification that reflects both the composition of the team and the hierarchical organization. Under plausible assumptions presented in the next section, results from individual time trials, and in particular the Prologue, enable us to identify the team composition effect and hence derive a measure of help intensity within team.

3.2 Methodology

The empirical strategy is broken down into two steps. In the first step, we identify the help intensity for each team and each Tour using data at the rider's level, and in particular the velocity of riders i) in the final classification of the Tour de France and ii) in the Prologue. In the second step, we use features of the distribution of help intensity to explain the evolution of overall performance inequality over time.

3.2.1 First step: Identification of helping intensity

In contrast to the setting of our model where teams are composed of one leader and one helper, in the Tour de France, each team is composed of 9 riders. Hence, a leader potentially receives the help of eight helpers.[19] Using

[19] From 1970 up until 1985, there were 10 riders in each team at the start of the Tour except for 1972 and 1973 where each team was composed of 11 riders. Since 1986, there

the terminology of our model, a leader l has velocity given by $v(z_l) = z_l + \sum_{i_l=1}^{8} s_{i_l} f(z_{i_l})$ where i_l indexes the helpers of leader l. Similarly, the velocity of each helper i_l is given as: $v(z_{i_l}) = z_{i_l} - a s_{i_l} z_{i_l}$. The average velocity of the helpers of leader l is then given as: $\overline{v}_l = \frac{1}{8}\sum_{i_l=1}^{8} v(z_{i_l}) = \overline{z}_l - \frac{a}{8}\sum_{i_l=1}^{8} s_{i_l} z_{i_l}$ where $\overline{z}_l = \frac{1}{8}\sum_{i_l=1}^{8} z_{i_l}$ is the average ability of the helpers of leader l.

Let $r_l \equiv v(z_l) - \overline{v}_l$ measure the within team velocity inequality in leader's l team. Note that $r_l = c_l + h_l$ where $c_l \equiv [z_l - \overline{z}_l]$ and $h_l \equiv \left[\sum_{i_l=1}^{8} s_{i_l}\left(\frac{a}{8} z_{i_l} + f(z_{i_l})\right)\right]$. This measure of within team inequality is decomposed into two terms: the team composition effect[20] c_l that captures the difference in ability between leader l and his "average" helper and the help intensity effect h_l that captures the extent of hierarchical organization within team.

Unfortunately, neither c_l nor h_l are directly measured in the data. However, as argued in Condition 4 above, during the prologue, $s_{i_l} = 0$ for all i_l and all l. Denoting z_i^p the ability of rider i at the Prologue and $v_i^p = v(z_i^p)$ his

are 9 riders by team.

[20] Broadly speaking, there are four "ability types" of riders participating at the Tour de France. Besides leaders that perform very well allround, there are sprinters that are (very) good at short time trials but (very) bad at mountain stages, "rouleurs" that are good allround and climbers that are very good at mountain stages but mediocre at (short) time trials. This means that swapping one ability type for another within team will generally have consequences both for the inequality in the final classification *AND* the Prologue. We exploit this relationship below in the identification of help intensity. In the robustness checks we also perform the same analysis but excluding sprinters (that are relatively easy to identify in the data) to further account for the composition effect.

velocity, this means that $v(z_i^p) = z_i^p$ for all i. It follows that the within team inequality at the Prologue reads as $r_l^p = z_l^p - \overline{z}_l^p \equiv c_l^p$ for a team l. Since the ability required to perform well at the Prologue might only be a subset of the abilities required to perform well in the final classification, rather than using r_l^p as a proxy for c_l and deriving h_l as $r_l - c_l^p$, we assume that c_l is linearly correlated with $c_l^p \equiv r_l^p$ and, h_l is orthogonal to r_l^p. We therefore identify h_l as the residuals of an orthogonal projection of r_l onto r_l^p. Indexing time by t, we have:

$$r_{lt} = \xi_0 + \xi_1 r_{lt}^p + e_{lt}. \tag{3}$$

By construction, the residuals e_{lt} of this equation are orthogonal to $r_{lt}^p \equiv z_{lt}^p - \overline{z}_{lt}^p$. For each team l in every Tour t we identify the help intensity h_{lt} as e_{lt}.

3.2.2 Second step: Estimation of the effect of help intensity on overall performance inequality

Let R_t be a measure of the overall performance inequality during Tour t. We use two measures of help intensity to explain the evolution of R_t over time. The first measure is the average help intensity at Tour t, say $e_t = \frac{1}{N_t} \sum_l e_{lt}$

where N_t is the number of teams at Tour t. The second measure is the range in the help intensity defined as $r_t^e = \max_l e_{lt} - \min_l e_{lt}$. We then consider the following equation:

$$R_t = \alpha_0 + \alpha_1 e_t + \alpha_2 r_t^e + \alpha_3' X_t + u_t,$$

where R_t is either proxied by the range, the reduced range, the lower range or the upper range in the final classification depending on the specification of the model and X_t are control variables. u_t are i.i.d. residuals.

The parameters α_1 and α_2 are the parameters of interest that relate the evolution of the distribution of help intensity (mean and range across teams) to the overall inequality. We also consider the following control variables:

1. 4 measures of internationalization:

    (a) The percentage of riders from the core[21] countries (1-globalization),

    (b) the percentage of riders from France (%French),

    (c) the percentage of riders from Italy (%Italian) and,

    (d) the percentage of riders from Spain (%Spanish),

[21] Belgium, the Netherlands, Luxembourg and Switzerland.

2. a time trend that captures technological development in a wide sense: let it be the type of bicycle and gear used, training methods, nutrition (including doping) etc.,

3. the difficulty of a particular Tour: we use as proxy the failing rate, i.e. the percentage of riders finishing.[22]

3.3 Data

For the empirical exercise, our main source of data is from http://www.tour-giro-vuelta.net/, a website managed by Michiel van Lonkhuyzen and data from http://www.letour.fr/HISTO/fr/TDF/. To correct for eventual mistakes and/or omission (a few distances and winning times), we cross checked between these two datasets but also with additional sources and in particular with Wikipedia for the total distance and the winning time and, http://www.ledicodutour.com/ and http://www.memoire-du-cyclisme.net for the general classification of the tour de France. Our database covers the Tour de France between 1947 and 2011.

For each participant appearing in the final classification of any Tour,

[22] We have also experimented with the direction of the tour de France (Pyrenees before or after the Alps) and the percentage of riders participating for the first time but this does not change the results.

we calculate two velocity measures: i) the *Tour velocity,* defined as the total distance of the tour divided by the participant's finishing time, for the period 1947-2011 and ii) the *Prologue velocity* defined as the distance of the Prologue divided by the participant's finishing time at the Prologue, available since the first Prologue in 1970 except for the Tours without Prologue, i.e. 1971, 1979, 1986, 1988, 2008 and 2011.

Overall Inequality

Our data enable us to derive a distribution of the Tour velocity for each Tour as well as its associated measures of inequality. In particular, we consider the range (velocity of the winner - velocity of the last rider) that constitutes an efficient estimate of inequality at time t (Parkinson, 1980).[23] For robustness purposes we also consider the range at specific parts of the distribution. The reduced range, defined as the difference between the inequality of the top and bottom 5% riders, is considered in order to control for extreme behavior: the high inequality observed in some years could be due to exceptional cluster of gifted riders such as the couple Hinault and Lemond in 80's or Armstrong and Ulrich in the early 2000's. Similarly, asymmetric ranges (upper and lower) are introduced to disentangle the factors affecting

---

[23] More precisely he showed that $\frac{R_t^2}{4\ln 2}$ converges to the spot volatility at time t.

the leaders and the helpers. These fours variables (range, restricted range, upper range and lower range) constitute the dependent inequality variables to be explained.

Figures 4 shows the evolution of the *Tour velocity* distribution over time. The figure clearly indicates a surge in inequality, represented here by the range of the *Tour velocity* distribution. This movement appears at the end of the 1960s. It is noticeable that this inflection turns out to be synchronous to the authorization of trade teams to participate and the increase in the prize money distributed.

To have a better insight of this remarkable development, we have also represented in Figure 5 the evolution of the *Tour velocity* density over time and in Figure 6 the evolution of the cumulative distribution of *Tour velocity*. Interestingly enough, it appears in Figure 5 that the higher *Tour velocity* inequality takes the form of a progressive modification of the shape of the distribution. While it was unimodal for the 1950s and 1960s, it progressively moves to a bimodal shape from the 1970s on. This indicates that a small group of top riders have improved their *Tour velocity* (relative to the contemporaneous mean *Tour velocity*) while the bulk of the riders have seen their performance deteriorate. Hence the inequality within the peloton is rising but in a very peculiar way. Figure 6 confirms this[24] but also informs us about the proportion of riders that have increased their performance relative to the contemporaneous mean overall velocity. This part can be found as the fixed point of the CDF and roughly corresponds to $0.6 - 0.7$. This means that $30 - 40\%$ of the riders improved their performances relative to the contemporaneous mean.

Figures 5 and 6 clearly indicate that inequality has strongly increased in the peloton of the Tour de France since the end of 1960s. Our intuition is that the modification of the organization of teams, for the reasons listed in Condition 3 above, impacted positively the inequality.

Within team inequality

We identify the leader of each team as the rider with the highest *Tour velocity*. Within each team, the remaining riders that finish the Tour are then considered as the helpers. The within team inequality between the leader and his helpers is given as $r_l = v(z_l) - \frac{1}{N_l - 1} \sum_{i_l=1}^{N_l-1} v(z_{i_l})$ where N_l is the number riders of team l that finished the Tour.[25]

[24] The twist of the cumulative distribution over time indicates the movement towards bimodality.

[25] Since 1970, on average, 30% of the riders does not finish the Tour. For the vast majority of these riders, the reason for not finishing is either a fall or sickness. We therefore herewith make the implicit assumption that the riders that do not finish the Tour were randomly drawn from the distribution of riders. Note however, that we control for the failing rate in our empirical analyses below.

Figure 7 shows the evolution of both the average within team inequality in the final classification, i.e. $\overline{r}_t = \frac{1}{N_t}\sum_{l=1}^{N_t} r_{lt}$, and at the Prologue $\overline{r}_t^p = \frac{1}{N_t}\sum_{l=1}^{N_t} r_{lt}^p$. Strikingly, the within team inequality at the Prologue is fairly stable over time. This supports the idea that changes in the composition of teams over time, if they actually have occurred, have not affected the inequality within team. In contrast, the within team inequality in the final classification follows the same general pattern as the overall inequality in the final classification. These two pieces of information together support the idea that the within team inequality is primarily driven by the rise in help intensity within team.

3.4 Results

We first present in Table 1 the results of an Ordinary Least Squares (OLS) estimation of Equation 3. The table indicates that the relationship between the within team inequality in the final classification and the within team inequality at the Prologue is positive and significant at 5%. The magnitude of the coefficient is economically very important. A 1km/h higher inequality at the Prologue is associated with a 0.03km/h higher inequality in the final classification. Stated otherwise, a leader that is 14 seconds faster at the

Prologue than his “average” helper is 4 minutes and 12 seconds faster in the final classification, ceteris paribus (without any organization within team).[26]

Figure 8 clearly shows that the average help intensity has increased over time following a similar pattern as the average within team inequality in the final classification. As shown in Table 2, simple OLS regressions of the average help intensity on the (deflated) money prices allocated to the winner of the Tour or its share in total money prices distributed during the Tour, clearly indicate that our measure of help intensity is significantly related with variables that are, according to our model's predictions, linked to the incentives to organize hierarchically within teams. In fact, these two variables alone explain about 57% and 15% respectively of the variance in our measure of help intensity over time.

Second, we present the results of OLS regressions of measures of overall performance inequaltiy on features of the distribution of help intensity in Table 3. The R-squared providing information on the quality of the regressions are extremely high (between 70% and 90%) for all measures of overall inequality. Our set of variables seems to constitute an adequate space to

[26] A typical Prologue is 10km long and covered in about 12 minutes (50km/h). A typical Tour de France is about 3,500km long and covered at an average velocity of about 39km/h.

analyze the inequality.[27] More importantly, the signs of the estimators are in line with our theoretical model. For all measures of overall inequality considered, our measure of help intensity has a positive impact, significant at 1%. This result provides a strong support for hierarchical organization as an explanation for the rise in performance inequality among riders in the Tour de France. Another remarkable result is that not only the sign but also the magnitude of the elasticity is robust to our choice of inequality measure,[28] ranging from 0.83 to 1.03. Our estimates indicate that an increase of help intensity leading to an increase of 1km/h in the velocity of a leader relative to that of his (average) helper leads to an increase of about 1km/h in the overall inequality.

Regarding the other candidates, we notice that only the between team inequality in help intensity is (weakly) significant and with the correct sign. All other variables are statistically insignificant. To summarize our results, it appears that hierarchical organization is the key variable that explains the

[27] Note that misspecification tests for autocorrelation (LM of Godfrey, 1978), heteroscedasticity (Breusch-Pagan, 1979), normality (Jarque and Bera, 1980) and structural break (Chow test with an unknown break date à la Andrews, 1993) are performed and support the idea of a correct specification. Finally, the presence of a unit root has been tested for each endogenous variable and rejected. All the tests are available upon request from the authors.

[28] It is noticeable that a simple t-test would lead to not reject the equality between these coefficients.

rise in productivity inequality in the Tour de France.

3.5 Robustness Check

To assess the robustness of our findings, two types of analyses are performed. First, we investigate further the effect of team composition on our results. Second, we confront another prediction of the model with an additional stylized fact of the evolution of the distribution of performance.

3.5.1 Stability of the relationship between hierarchical organization and overall performance inequality

We consider controlling further for the team composition effect by excluding sprinters from the analysis. To flag sprinters, we collected additional information about the classification of riders during a stage that ended with a massive sprint. We then labeled "sprinter" every rider that i) finished within the 20 best riders during the sprint and ii) did not finish within the 20 best riders in the final classification. Having flagged sprinters, we run our two-step procedure excluding sprinters.

The results of the first and second step regressions are reported in Table 4. The table clearly indicates that both the coefficients of the first and second step are of similar magnitude to the one presented in Table 3. This means that controlling further for the team composition effect by excluding sprinters from the sample does not affect the results.

3.5.2 Can the model reproduce the stylized facts?

To further evaluate the empirical prediction of the model, we propose the following test. Consider N teams and suppose that each team has only one leader and all 9 riders of each team finish the tour. We also assume that initially, riders are assigned at random to teams and help intensity is zero for every riders (autarky). This means that the final classification reflects the true distribution of ability. In particular, the velocity of the N^{th} rider relative to that of the $N+1^{th}$ rider reflects their ability differential. Following, for instance, a convexification of the reward function, suppose that the new equilibrium exhibits a strict stratification of riders: all N leaders are strictly better than any of the $8 \times N$ helpers. In the final classification, the first riders are the leaders of the various teams and the last $8 \times N$ riders are their helpers (or riders riding individually). The performance of all leaders increases while the performance of all helpers decreases holding everything else constant. The model has three important predictions. First, the within team inequality increases in all teams. Second, the overall inequality increases too. Third and most importantly, the performance of the N^{th} rider (the least able leader) increases while the performance of the $N+1^{th}$ rider (the best helper) decreases. This means that at constant distribution of ability, we should observe a movement of riders above the $100 \times \left(1 - \frac{N}{9N}\right) = 100 \times \frac{8}{9} \approx 90^{th}$ quantile away from riders below the 90^{th} quantile.

Figure 6 clearly shows that the distribution becomes more unequal over time but it is striking to see that all curves seem to be twisting clock wise with a twisting point at the 60^{th} quantile. Although this sketchy model predictes a twist at the 90^{th} quantile, one should bear in mind that 1) we have assumed a strict stratification which is only a special case in the economy depicted in Section 2 and 2) the model depicts economies without performance shocks (no sickness during the tour, no falls, no exclusion for doping etc.) and with perfect information about the ability for all riders. With this in mind, we take the results presented in Figure 6 as supporting our hypothesis that the increase in the performance inequality is primarily due to an increase in the hierarchical organization of teams (via help intensity).

4 Discussion

This paper investigates the relationship between hierarchical organization and performance inequality within and between organizations. An equilibrium theory of the organization of work in an economy with an implicit market for productive time is first presented. In this economy, agents have limited productive time and can choose to produce in autarky, buy productive time from helpers to increase own production or, sell their productive time to a leader and thereby give up own production. This implicit market gives rise to the formation of teams, organized in hierarchies with one leader at the top and helpers below. We prove that an equilibrium exists and is efficient and show that relative to autarky, hierarchical organization leads to higher within and between team payoffs/productivity inequality.

To illustrate the main prediction of our theoretical model, i.e. team organization increases performance inequality, we propose an empirical analysis in the context of professional road cycling. Considering such a framework is novel in this literature and has several key advantages compared to other markets to study the existence of an implicit market for productive time. These three key advantages are i) road cycling exhibits a clear change in the incentives to organize work within team since the end of the 1960's, ii) a direct measure of individual productivity is available in that sector via riders' velocity and iii) results from individual time trials enable us to identify the composition of team and hence derive a measure of help intensity. Results of performance inequality regressions robustly show that leaders' velocity increased significantly (economically and statistically) more than that of helpers because of the increasing help intensity within teams. This supports hence the model's prediction of a positive relationship between hierarchical organization and productivity inequality.

Although the core model developed in the paper relies on some of the characteristics of professional road cycling, we believe this core model constitutes an adequate corpus to analyze the relationship between earnings inequality and hierarchical organization of firms in many other industries. Several extensions may be proposed in the future to make it fit with other industries. First, in the core theoretical model presented in this paper, the size of teams is fixed exogenously following hence the rule that applies in professional road cycling. It is obvious that such an hypothesis does not hold in other organizations. For example a law firm can engage as many collaborators as possible to help a leading lawyer prepare for a trial. Law firms can to some extent adjust not only the intensive but also the extensive margin of organizational structure. Second, the core model considers a single event (the Tour de France). Though the Tour de France is the single most important event of the season, teams usually engage in many other events (such as the classics and shorter tours). The hierarchical organization of a team may thus vary in function of the race: helpers during major races may be leaders during minor races and vice versa. Interestingly enough, this dynamic organization can be perceived as part of the compensation offered by the leaders (during major events) to their helpers. Modelling this dynamic organization of team seems particularly interesting. Third, the core model presented in this paper excludes any role for team managers. One may think of several interesting ways of including managers into the model. One way would be to assume that managers improve teams' performance. Another, perhaps more interesting, extension would be to incorporate the strategic role played by managers when performance outcome is uncertain. Finally, applying the model to segments of the labor market raises the question of the interpretation of the pricing function $p(.)$. These various extensions do not minimize the importance of the core model presented in the paper, they rather highlight its importance in this literature.

References

Acemoglu, D. (2003): "Cross-Country Inequality Trends," *Economic Journal*, 113, 121–49.

Acemoglu, D., and D. Autor (2010): "Skills, Tasks and Technologies: Implications for Employment and Earnings," NBER Working Papers 16082, National Bureau of Economic Research, Inc.

Andrews, D. (1993): "Tests for Parameter Instability and Structural Change with Unknown Change Point," *Econometrica*, 61(4), 821–56.

Becker, G. S. (1973): "A Theory of Marriage: Part I," *Journal of Political Economy*, 81(4), 813–46.

Breusch, T. S., and A. R. Pagan (1979): "A Simple Test for Heteroscedasticity and Random Coefficient Variation," *Econometrica*, 47(5), 1287–94.

Chiappori, P.-A., A. Galichon, and B. Salanié (2011): "The Roomate Problem with Transferable Utilities or Marriage with Same-Sex Unions," mimeo.

Chiappori, P.-A., R. McCann, and L. Nesheim (2010): "Hedonic price equilibria, stable matching, and optimal transport: equivalence, topology, and uniqueness," *Economic Theory*, 42(2), 317–354.

Edmans, A., X. Gabaix, and A. Landier (2009): "A Multiplicative Model of Optimal CEO Incentives in Market Equilibrium," *Review of Financial Studies*, 22(12), 4881–4917.

Fox, J. T. (2009): "Firm-Size Wage Gaps, Job Responsibility, and Hierarchical Matching," *Journal of Labor Economics*, 27(1), 83–126.

Gabaix, X., and A. Landier (2008): "Why Has CEO Pay Increased So Much?," *The Quarterly Journal of Economics*, 123(1), 49–100.

Garicano, L., and T. Hubbard (2009): "Earnings Inequality and Coordination Costs: Evidence From U.S. Law Firms," NBER Working Papers 14741, National Bureau of Economic Research, Inc.

Garicano, L., and T. N. Hubbard (2005): "Hierarchical sorting and learning costs: Theory and evidence from the law," *Journal of Economic Behavior & Organization*, 58(2), 349–369.

Garicano, L., and E. Rossi-Hansberg (2004): "Inequality and the Organization of Knowledge," *American Economic Review*, 94(2), 197–202.

——— (2006): "Organization and Inequality in a Knowledge Economy," *The Quarterly Journal of Economics*, 121(4), 1383–1435.

Godfrey, L. G. (1978): "Testing for Higher Order Serial Correlation in Regression Equations When the Regressors Include Lagged Dependent Variables," *Econometrica*, 46(6), 1303–10.

Jarque, C. M., and A. K. Bera (1980): "Efficient tests for normality, homoscedasticity and serial independence of regression residuals," *Economics Letters*, 6(3), 255–59.

Lucas, R. E. (1978): "On the Size Distribution of Business Firms," *Bell Journal of Economics*, 9(2), 508–523.

McGann, B., and C. McGann (2006): *The Story of the Tour de France: volume 1 and 2*. Dog Ear Publishing.

Parkinson, M. (1980): "The extreme value method for estimating the variance of the rate of return," *Journal of Business*, 53(1), 65–65.

Rajan, R. G., and J. Wulf (2006): "The Flattening Firm: Evidence from Panel Data on the Changing Nature of Corporate Hierarchies," *Review of Economics and Statistics*, 88(4), 759–773.

Rosen, S. (1974): "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition," *Journal of Political Economy*, 82, 34–55.

——— (1982): "Authority, Control, and the Distribution of Earnings," *Bell Journal of Economics*, 13(2), 311–323.

Rubinstein, J. S., D. E. Meyer, and J. E. Evans (2001): "Executive Control of Cognitive Processes in Task Switching," *Journal of Experimental Psychology: Human Perception and Performance*, 27(4), 763–797.

Sattinger, M. (1993): "Assignment Models of the Distribution of Earnings," *Journal of Economic Literature*, 31(2), 831–80.

Tervio, M. (2008): "The Difference That CEOs Make: An Assignment Model Approach," *American Economic Review*, 98(3), 642–68.

Tinbergen, J. (1956): "On the Theory of Income Distribution," *Weltwirtschaftliches Archiv*, 77(1), 156–75.

Villani, C. (2009): *Optimal Transport, Old and New*. Springer-Verlag Berlin Heidelberg.

Appendix A: Second order conditions:

The second order condition to the leader's problem reads as:

$$\begin{aligned} w_h''(z_h) &> p''(z_l+f(z_h))\left(f'(z_h)\right)^2+p'(z_l+f(z_h))f''(z_h)+(1-a)^2p''(z_h(1-a)) \\ \text{if } 0 &< p(z_l+f(z_h))-p(z_l)-p(z_h)+p(z_h(1-a)) \\ & \quad and \\ w_0''(z_h) &> p''(z_h)>0 \text{ otherwise.} \end{aligned} \tag{4}$$

Similarly, the second order condition to the helper's problem reads as:

$$\begin{aligned} w_l''(z_l) &> p''(z_l+f(z_h)) \\ \text{if } v(z_l+f(z_h))-p(z_l) &> p(z_h)-p(z_h(1-a)) \\ & \quad and \\ w_0''(z_l)-p''(z_l) &> 0 \text{ otherwise.} \end{aligned} \tag{5}$$

An important remark is that our standing assumptions SAI and SAII imply that $w_i''(z_i)>0$ for $i=l,h$ so that equilibrium payoffs functions are convex.

Appendix B: Proofs of propositions

We first prove Lemma 13 that will be used in the proof of Proposition 3.

Lemma 13 *Under SA II, we have* $p(x_2 + \Delta_2) - p(x_2) \geq p(x_1 + \Delta_1) - p(x_1)$ *for all* $x_2 \geq x_1 \geq 0$ *and* $\Delta_2 \geq \Delta_1 \geq 0$.

Proof. Since from SA II.2, $p'(z) > 0$ for all $z \geq 0$, it follows that $p(x_2+\Delta_2) \geq p(x_2 + \Delta_1)$ for all $\Delta_2 \geq \Delta_1 \geq 0$.

It remains to show that $p(x_2 + \Delta_1) - p(x_2) \geq p(x_1 + \Delta_1) - p(x_1)$ for all $x_2 \geq x_1 \geq 0$ and $\Delta_1 \geq 0$.

Write $g_\Delta(x) \equiv p(x + \Delta) - p(x)$ with $\Delta \geq 0$. By definition we have $g'_\Delta(x) = p'(x + \Delta) - p'(x)$.

Since $p'' \geq 0$ from SA II.3, p' is increasing over x so that $g'_\Delta(x) = p'(x + \Delta) - p'(x) \geq 0$. It follows that $g_{\Delta_1}(x_2) = p(x_2 + \Delta_1) - p(x_2) \geq p(x_1 + \Delta_1) - p(x_1) = g_{\Delta_1}(x_1)$ for $x_2 \geq x_1$. ∎

Proof of proposition 3: More able riders become leaders.

Proof. Take a team of riders with respective ability x and y with $x \geq y$ without loss of generality. This team's surplus is $Y(x, y) \equiv \max_s p(x + f(y)s) + p(y(1 - as))$ when x is the leader and $Y(y, x) \equiv \max_s p(y + f(x)s) +$

$p(x(1-as))$ when y is the leader. To prove that x will always be the leader we need to prove that $Y(x,y) \geq Y(y,x)$. Denote $s^0 = s^*(y,x) = \arg\max_s p(y + f(x)s) + p(x(1-as))$ and denote $s^1 = s^*(x,y) = \arg\max_s p(x + f(y)s) + p(y(1-as))$. By definition we have:

$$Y(x,y) = p(x + f(y)s^1) + p(y(1-as^1)) \geq p(x + f(y)s^0) + p(y(1-as^0)).$$

Hence, it is enough to prove that $p(x + f(y)s^0) + p(y(1-as^0)) > p(y + f(x)s^0) + p(x(1-as^0))$ for all $1 \geq s^0 \geq 0$. Rearranging terms, we aim at proving that the following inequality holds for all s^0 and $x \geq y$:

$$p(x + f(y)s^0) - p(y + f(x)s^0) \geq p(x(1-as^0)) - p(y(1-as^0)). \quad (6)$$

Write $x_1 = y(1-as^0)$ and $x_1 + \Delta_1 = x(1-as^0)$ where $\Delta_1 = x - y + as^0(y-x)$ and $x_2 = y + f(x)s^0$ and $x_2 + \Delta_2 = x + f(y)s^0$ where $\Delta_2 = x - y + s^0(f(y) - f(x))$. Note that $x_2 \geq x_1$ for all $y \in Z$ and $s^0 \in [0,1]$. Inequality 6 can be written as:

$$p(x_2 + \Delta_2) - p(x_2) \geq p(x_1 + \Delta_1) - p(x_1) \text{ with } x_2 \geq x_1.$$

Following Lemma 13, a sufficient condition for this inequality to hold is $\Delta_2 \geq \Delta_1$. Replacing Δ_1 and Δ_2 by their expression in terms of x and y and rearranging yields:

$$\Delta_2 \geq \Delta_1 \Leftrightarrow a(x-y) \geq f(x) - f(y).$$

Hence, from SA I.2 we have $p(x+f(y)s^0)+p(y(1-as^0)) \geq p(y+f(x)s^0)+p(x(1-as^0))$ for all s^0 and $x \geq y$. This means that $Y(x,y) \geq Y(y,x)$ for all $x \geq y$. The surplus of a team is therefore always higher when the most able rider is helped by the least able one. ■

Proof of Proposition 4: conditions for a corner solution of $s^*(x,y)$.

We first prove Lemma 14 that will be used to prove Proposition 4.

Lemma 14 *Under SA I and SA II, for all feasible teams* $(z_l, z_h) \in Z^2$*, the surplus function* $Y(z_l, z_h, s)$ *is a strictly convex function of helping time on* $s \in [0,1]$.

Proof. Take a team of riders with respective ability x and y and with $x \geq y$ without loss of generality. From Proposition 3, rider x becomes the leader and rider y the helper. This team's surplus is therefore $Y(x,y,s) =$

$p(x + f(y)s) + p(y(1 - as))$. The slope of the surplus with respect to helping time obtains as:

$$\frac{\partial Y(x,y,s)}{\partial s} = p'(x + f(y)s)f(y) - p'(y(1 - as))ay.$$

The curvature of the surplus with respect to helping time is given by:

$$\frac{\partial^2 Y(x,y,s)}{\partial s^2} = p''(x + f(y)s)\,(f(y))^2 + p''(y(1 - as))\,(ay)^2\,.$$

From SA I.1 and SA II.3, we have $\frac{\partial^2 y(x,y,s)}{\partial s^2} > 0$. The surplus function $Y(x,y,s)$ is strictly convex on $s \in [0,1]$. ■

We can now prove Proposition 4.

Proof. Take a team of riders with respective ability x and y with $x \geq y$ without loss of generality. From proposition 3, rider x becomes the leader and rider y the helper. This team has surplus equal to $Y(x,y,s) = p(x + f(y)s) + p(y(1 - as))$. From Lemma 13, we know that $Y(x,y,s)$ is strictly convex in s on $s \in [0,1]$ for all x and y. This means that:

$$s^*(x,y) = \begin{cases} 1 \text{ iff } Y(x,y,1) > Y(x,y,0) \\ 0 \text{ otherwise} \end{cases}.$$

Using the definition of Y and rearranging yields:

$$s^*(x,y)=\begin{cases}1 \text{ iff } p(x+f(y))-p(x)>p(y)-p(y(1-a))\\ \quad\quad\quad\quad 0 \text{ otherwise}\end{cases}.$$

■

Proof of Proposition 7: Under SA I and SA II, in equilibrium, more able leaders are matched with more able helpers.

Proof. By the implicit function theorem, write $z_h = z_h(z_l)$ the solution of Equation 1 and $z_l = z_l(z_h)$ the solution of Equation 2.

First, suppose that $z_h(z_l)$ and $z_l(z_h)$ are differentiable. Then, totally differentiating Equation 1 with respect to z_l and Equation 2 with respect to z_h and rearranging yields:

$$z_h'(z_l) = \frac{p^{''}(z_l+f(z_h))f'(z_h)}{w_h''(z_h)-p''(z_l+f(z_h))\left(f'(z_h)\right)^2+p'(z_l+f(z_h))f''(z_h)+(1-a)^2p''(z_h(1-a))},$$

$$z_h'(z_l) = \frac{p^{''}\left(z_l+f(z_h)\right)f'(z_h)}{w_l''(z_l)-p''(z_l+f(z_h))}.$$

From the second order conditions in Equations 4 and 5, the denominators are strictly positive. Hence, $z_h'(z_l))>0$ since $p^{''}(z_l+f(z_h(z_l)))f'(z_h(z_l))>0$ from SA I.3 and SA II.3 and $z_l'(z_h))>0$ since $p^{''}(z_l(z_h)+f(z_h))f'(z_h)>0$

from SA I.3 and SA II.3.

Suppose now that $z_h(z_l)$ and $z_l(z_h)$ are not differentiable. We can still prove that in equilibrium, more able leaders get more able helpers. Take two teams that arise in equilibrium say (x_i, y_i), $i = 0, 1$ where x_i is the ability of the leader and y_i the ability her helper. Without loss of generality, suppose that $x_1 = x_0 + h$ with $h > 0$. From the second order conditions in Equations 4 and 5, we know that the wage profiles are steeper than the productivity profiles. Formally, and using the helper's problem for instance, we have that:

$$\lim_{h \to 0} \frac{w_l'(x_0 + h) - w_l'(x_0)}{h} = w_l^{''}(x_0) > p^{''}(x_0) = \lim_{h \to 0} \frac{p'(x_0 + h + f(y_0)) - p'(x_0 + f(y_0))}{h}.$$

Using the first order condition in Equation 2 to replace $w_l'(.)$ obtains:

$$\begin{aligned} \lim_{h \to 0} \frac{p'(x_0 + h + f(y_1)) - p'(x_0 + f(y_0))}{h} &> \lim_{h \to 0} \frac{p'(x_0 + h + f(y_0)) - p'(x_0 + f(y_0))}{h} \\ &\Leftrightarrow \\ y_1 &> y_0. \end{aligned}$$

It follows that in equilibrium, more able leaders are matched with more able helpers. ■

Proof of Proposition 8: Under SA I and SA II, in equilibrium, no riders of ability lower than $z^{(l)}$ become leaders and no riders of ability higher than $z^{(l)}$ rider on their own.

Proof. Take any rider of ability $z \in Z$. The payoffs of this rider are $w_l(z)$ as a leader, $w_h(z)$ as a helper and $w_0(z) \equiv p(z)$ as an individual rider. A payoffs maximizing rider will therefore choose the role leading to $W(z) = \max\{w_l(z), w_h(z), p(z)\}$. We are looking for the upper envelop $W(z)$ of the graph of payoffs $\{w_l(z), w_h(z), p(z)\}$ in z.

Without further restrictions, we already know from the first order conditions that leaders' payoffs function $w_l(z)$ is strictly steeper than that of individual riders $p(z)$, i.e. $w_l'(z) = p'(z + f(z_h)) > p'(z)$ from SA II. Let $z^{(l)}$ be the ability of riders so that $w_l(z^{(l)}) = p(z^{(l)})$. This implies that $w_l(z) < p(z) \leq W(z)$ for all $z < z^{(l)}$ and $W(z) \geq w_l(z) > p(z)$ for all $z > z^{(l)}$. It follows that in equilibrium, there are no leaders of ability lower than $z^{(l)}$ and no individual riders of ability higher than $z^{(l)}$. Stated otherwise, riders of ability lower than $z^{(l)}$ either become a helper or an individual rider, i.e. $w_l(z) < W(z)$ for all $z < z^{(l)}$, while riders of ability higher than $z^{(l)}$ either become a helper or a leader, i.e. $w_0(z) < W(z)$ for all $z > z^{(l)}$. ■

Proof of Proposition 9: Under SA I and SA II, an optimal assignment (L, H, γ) exists.

Proof. Sketch of the proof.[29] Step 1: We follow Chiappori, Galichon and Salanie (2011) and show that the social planner problem SPP can be written as the primal program of a classical Monge-Kantorovich problem with symmetric surplus function. Step 2: We then derive the properties of our model from the properties of the Monge-Kantorovich problem that have been studied in Villani (2009).

Step 1: Consider a pair of riders (x, y) such that $d\gamma(x, y) + d\gamma(y, x) > 0$. Without loss of generality, suppose that $Y_1(x, y) > Y_1(y, x)$. The contribution of this pair to program (P1) would be largest when $d\gamma(x, y) > d\gamma(y, x) = 0$. This means that a solution γ for SPP is necessarily such that $d\gamma(z_l, z_h) \geq d\gamma(z_h, z_l) = 0$ whenever $Y_1(z_l, z_h) = \max(Y_1(z_l, z_h); Y_1(z_h, z_l))$. The contribution of a pair (z_l, z_h) so that $Y_1(z_l, z_h) = \max(Y_1(z_l, z_h); Y_1(z_h, z_l))$ is therefore

$$Y_1(z_l, z_h)d\gamma(z_l, z_h) + Y_1(z_h, z_l)d\gamma(z_h, z_l) = Y_1(z_l, z_h)d\gamma(z_l, z_h).$$

[29] An alternative proof of existence is to show that $\Gamma(\mu)$ is tight, hence compact by Prokhorov's Theorem. Since Y_1 is continuous, there exists a solution γ of (P1). This proof is available from the authors upon request.

Define $\widetilde{Y}_1(z_l, z_h) := \max(Y_1(z_l, z_h); Y_1(z_h, z_l))$ and let $d\widetilde{\gamma}(z_l, z_h) := d\widetilde{\gamma}(z_h, z_l) := \frac{d\gamma(z_l,z_h)+d\gamma(z_h,z_l)}{2}$. Obviously, since $\widetilde{Y}_1$ is symmetric, we have that:

$$\widetilde{Y}_1(z_l, z_h)d\widetilde{\gamma}(z_l, z_h)+\widetilde{Y}_1(z_h, z_l)d\widetilde{\gamma}(z_h, z_l) = Y_1(z_l, z_h)d\gamma(z_l, z_h)+Y_1(z_h, z_l)d\gamma(z_h, z_l).$$

Note that $\widetilde{\gamma}$, not only is symmetric, but also satisfies the feasibility constraint of Definition 5. For symmetric measures, these constraints can be re-written as:

$$\begin{aligned} \int_{z_l \in Z} d\widetilde{\gamma}(z_l, z) &= \frac{1}{2}d\mu(z) \text{ (a)} \\ \int_{z_h \in Z} d\widetilde{\gamma}(z, z_h) &= \frac{1}{2}d\mu(z) \text{ (b)} \\ d\widetilde{\gamma}(z_l, z_h) &= d\widetilde{\gamma}(z_h, z_l) \text{ (c).} \end{aligned}$$

Let $\widetilde{\Gamma}(\frac{1}{2}\mu, \frac{1}{2}\mu)$ be the set of measures $\widetilde{\gamma}$ satisfying constraints (a), (b) and (c). Program (P1) is therefore equivalent to program (P2) below:

$$SPP^* = \max_{(L,H)\in Z^2, \widetilde{\gamma}\in\widetilde{\Gamma}(\frac{1}{2}\mu,\frac{1}{2}\mu)} \left\{ \int_{Z\setminus L\cup H} p(z)d\mu(z) + \int_{L\times H} \widetilde{Y}_1(z_l, z_h)d\widetilde{\gamma}(z_l, z_h) \right\}. \tag{P2}$$

Define program (P2') as the same maximization as in program (P2) but without constraint (c) and let SPP^*_{MK} be its value. Note that program (P2') reads as the primal program of a Monge-Kantorovich transportation problem with symmetric surplus. Theorem 4.1 in Villani (2009) asserts that there exists a solution, say $\widetilde{\widetilde{\gamma}}$, to (P2') since $\widetilde{Y}_1$ is upper-semicontinuous. It is now easy to see that $\widetilde{\gamma}$ defined by $d\widetilde{\gamma}(z_l, z_h) := \frac{d\widetilde{\widetilde{\gamma}}(z_l,z_h)+d\widetilde{\widetilde{\gamma}}(z_h,z_l)}{2}$ is also a solution of (P2') and satisfies constraint (c) so that it is a solution of (P2). It follows that γ defined by[30] $d\gamma(z_l, z_h) := \begin{cases} 2d\widetilde{\gamma}(z_l,z_h) \text{ if } Y(z_l,z_h)=\widetilde{Y}_1(z_l,z_h) \\ 0 \text{ else} \end{cases}$ is a solution of (P1). We conclude that there exists a solution to (P1). As a by-product we have also shown that $SPP^* = SPP^*_{MK}$. This last result will be used in the proof of Proposition 10 below. ■

Proof of proposition 10: Under SA I and SA II, there is duality, i.e. $SPP^* = DP^*$.

Proof. Sketch of the proof: Step 1: In the proof of Proposition 9, following Chiappori, Galichon and Salanie (2011), we have shown that the social planner problem can be re-written as the primal program of a classical Monge-Kantorovich problem with symmetric surplus function. The value of the two programs are equal, i.e. $SPP^* = SPP^*_{MK}$. Step 2: We proceed in a similar fashion and show that the dual of the social planner program can be re-

[30] In case $Y_1(z_l, z_h) = Y_1(z_h, z_l) = \widetilde{Y}_1(z_l, z_h)$ the distribution of roles within teams does not matter. For all $\alpha \in [0, 1]$, $d\gamma(z_l, z_h) = \alpha \times 2d\widetilde{\gamma}(z_l, z_h)$ and $d\gamma(z_h, z_l) = (1 - \alpha) \times 2d\widetilde{\gamma}(z_l, z_h)$ is solution.

written as the dual of the associated Monge-Kantorovich problem, and show that the values of these two dual programs are also equal. Step 3: Since Villani (2009) asserts that there is duality in the associated Monge-Kantorovich problem, we conclude that there is duality in our one-sided assignment model.

Step 1: The social planner program.

As shown in the proof of Proposition 9, we have $SPP^* = SPP^*_{MK}$.

Step 2: The dual program.

The dual program associated to program (P2') introduced in the proof of proposition 9 reads as:

$$\begin{aligned} DP^*_{MK} &= \min_{w_l, w_h} \int_Z \left(\widetilde{w}_l(z) + \widetilde{w}_h(z)\right) \frac{1}{2} d\mu(z) \qquad \text{(D2')} \\ & s.t. \\ \widetilde{w}_i(z) &\geq p(z)\ \forall i = h, l \text{ and for all } z \in Z \text{(i')} \\ \widetilde{w}_l(z_l) + \widetilde{w}_h(z_h) &\geq \widetilde{Y}_1(z_l, z_h) \text{ for all } z_l, z_h \in Z^2 \text{ (ii')}. \end{aligned}$$

It is easy to see that (D1) is equivalent to (D2') but with the additional symmetry constraint $\widetilde{w}_l = \widetilde{w}_h$. We therefore have by definition $DP^* \geq DP^*_{MK}$. Let $(\widetilde{w}_l, \widetilde{w}_h)$ be a solution of (D2') which exists from Theorem 5.10 in Villani (2009). Obviously, (w_l, w_h) defined as $w_l := \frac{\widetilde{w}_l + \widetilde{w}_h}{2} := w_h$ is also a solution of (D2'). Since (w_l, w_h) satisfies the symmetry constraint it is also a solution of (D1). We therefore have: $DP^* = DP^*_{MK}$.

Step 3: Duality.

Since Theorem 5.10 part i) in Villani (2009) asserts that $SPP^*_{MK} = DP^*_{MK}$ as long as $\widetilde{Y}_1$ is upper-semicontinuous, we conclude that $SPP^* = DP^*$: there is duality in the one-sided assignment model. ■

Proof of proposition 11: A feasible tuple $((w_h(z), w_l(z)), (L, H, \gamma))$ that solves both the primal and dual program maximizes riders payoffs.

Proof. Suppose that $w(z) := \max\{w_h(z), w_l(z)\}$ solves the dual program (D1) and (L, H, γ) solves the primal program. We then have:

$$\begin{aligned}\int_Z \max(w_l, w_h) d\mu(z) &= \int_L w_l(z) d\mu(z) + \int_{H \setminus L \cap H} w_h(z) d\mu(z) + \int_{Z \setminus L \cup H} p(z) d\mu(z) \\ &= \int_{L \times H} Y_1(z_l, z_h) d\gamma(z_l, z_h) + \int_{Z \setminus L \cup H} p(z) d\mu(z),\end{aligned}$$

where the first equality follows since by definition $w_l(z) = w_h(z) = w(z)$ for $z \in L \cap H$, $w_l(z) = w(z)$ for $z \in L \setminus L \cap H$ and $w_h(z) = w(z)$ for $z \in H \setminus L \cap H$, and the second from duality $SPP^* = DP^*$ proved in Proposition 10.

Hence:

$$
\begin{aligned}
w_h(z_h) + w_l(z_l) &= Y_1(z_l, z_h) \text{ for } \gamma - a.e.\ (z_l, z_h) \in L \times H \\
&\text{and} \\
w_l(z) &= w_h(z) = p(z) \text{ for } \mu - a.e.\ z \in Z \backslash L \cup H.
\end{aligned}
$$

Take a leader $z_l^* \in L$ that is matched with a helper $z_h^* \in H$ in equilibrium (i.e. so that $d\gamma(z_l^*, z_h^*) > 0$). The riders of this team get respectively $w_h(z_h^*) = Y_1(z_l^*, z_h^*) - w_l(z_l^*)$ and $w_l(z_l^*) = Y_1(z_l^*, z_h^*) - w_h(z_h^*)$. Since $(w_h(z), w_l(z))$ solves the dual program, the feasibility constraints are satisfied so that $w_l(z_l^*) \geq Y_1(z_l^*, z_h) - w_h(z_h)$ for all $z_h \in Z$ and $w_h(z_h^*) \geq Y_1(z_l, z_h^*) - w_l(z_l)$ for all $z_l \in Z$. It follows that helper z_h^* maximizes the payoffs of leader z_l^* and leader z_l^* maximizes the payoffs of helper z_h^*. ■

Proof of corollary 12: The pair of payoffs functions (w_l, w_h) that maximizes riders' payoffs is Pareto Optimal.

Proof. A feasible tuple $((w_h(z), w_l(z)), (L, H, \gamma))$ that solves the primal and dual program maximizes riders' payoffs from Proposition 11. Since (L, H, γ) solves the primal program, the pair of equilibrium payoffs (w_l, w_h) maximizing riders payoffs also maximizes social welfare. ■

Table 1: Identification of help intensity: regression of within team inequality at Tour de France on within team inequality at the Prologue.

Variable	Within team inequality
Within team inequality Prologue	0.0258**
	(0.011)
Constant	0.3632***
	(0.023)
Observations	648
R-squared	0.008

Standard errors in parentheses
* p<0.10, ** p<0.05, *** p<0.01

Table 2: Help intensity and incentives to organize hierarchically.

Variable	Help Intensity	Help intensity
Price money winner	0.2265***	
	(0.034)	
Share winner		0.0072**
		(0.003)
Constant	-0.1291***	-0.1135***
	(0.020)	(0.040)
Observations	34	34
R-squared	0.574	0.149

Standard errors in parentheses
* p<0.10, ** p<0.05, *** p<0.01

Table 3: Regressions on the range, reduced range, lower range and upper range.

Variable	Range	Upper	Lower	Reduced
Help intensity	1.0287***	0.9003***	0.8287***	0.8645***
	(0.231)	(0.183)	(0.250)	(0.151)
Between team help intensity	0.4149*	0.6240***	-0.0927	0.2657*
	(0.238)	(0.188)	(0.257)	(0.155)
Trend	-0.0022	0.0046	-0.0010	0.0018
	(0.004)	(0.003)	(0.004)	(0.003)
% Core countries	-0.2543	-0.0597	-0.3697	-0.2147
	(0.397)	(0.314)	(0.430)	(0.260)
% French	0.0421	-0.1725	0.4932	0.1603
	(0.344)	(0.273)	(0.373)	(0.225)
% Spanish	-0.1834	-0.0238	0.1742	0.0752
	(0.457)	(0.362)	(0.495)	(0.299)
% Italian	-0.0822	0.0099	-0.0410	-0.0155
	(0.332)	(0.263)	(0.359)	(0.217)
% Failing rate	0.2948	-0.1761	0.1009	-0.0376
	(0.232)	(0.183)	(0.251)	(0.151)
Constant	4.8017	-8.4119	2.3388	-3.0366
	(8.001)	(6.333)	(8.668)	(5.235)
Observations	35	35	35	35
R-squared	0.830	0.935	0.673	0.914

Standard errors in parentheses
* $p<0.10$, ** $p<0.05$, *** $p<0.01$

Table 4: Regressions on the range, reduced range, lower range and upper range excluding sprinters.

First Step				
Within team inequality Prologue	0.0274**			
Second Step				
Variable	Range	Upper	Lower	Reduced
Help intensity	1.0812***	1.0034***	0.7970***	0.9002***
Between team help intensity	0.1900	0.5096**	-0.1158	0.1969

Standard errors in parentheses
* $p<0.10$, ** $p<0.05$, *** $p<0.01$

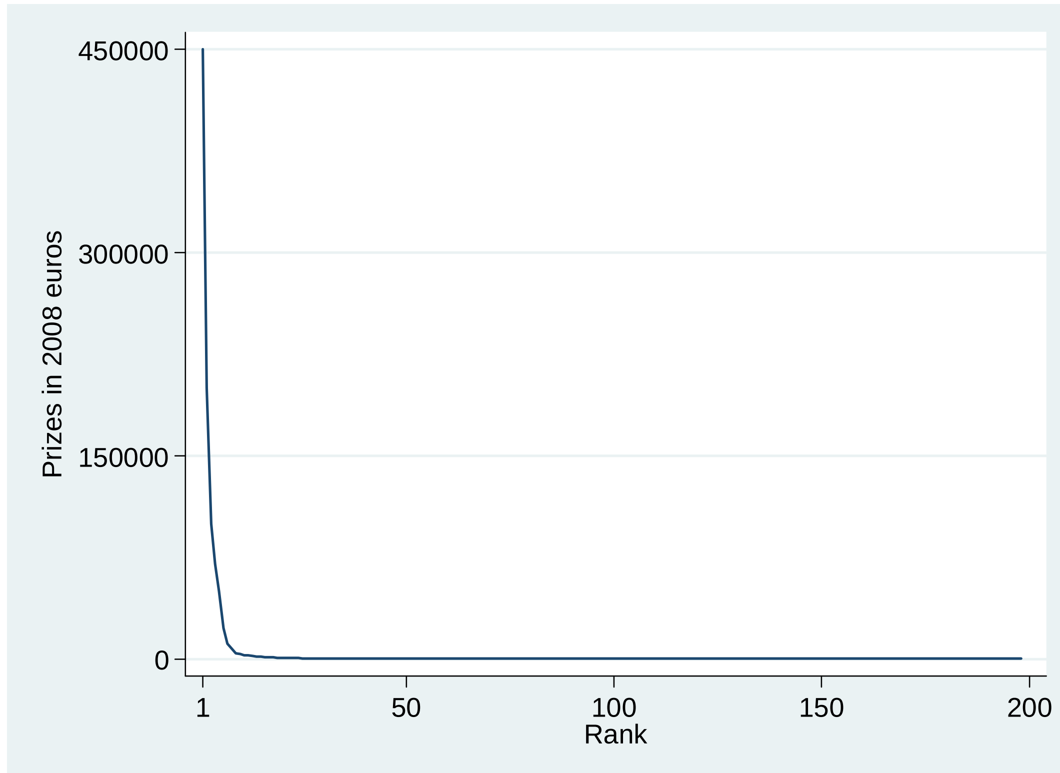

Figure 1: Distribution of prize money by rank in the final classification. Note: in 2008 euros.

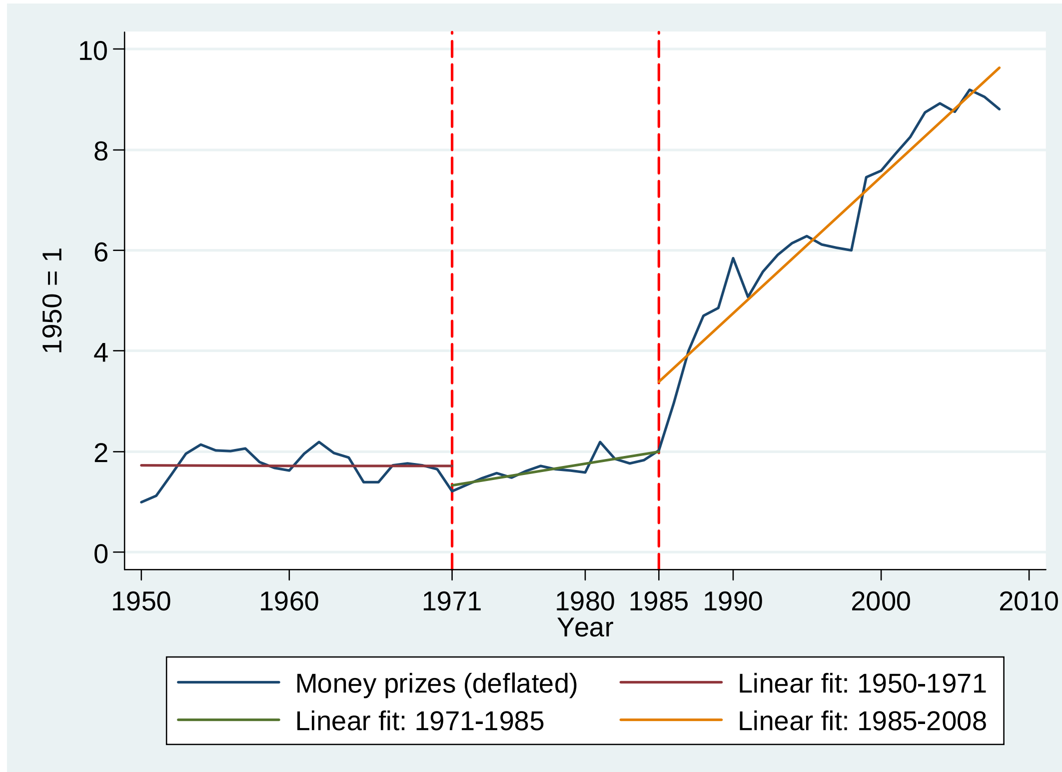

Figure 2: Evolution of total prize money distributed during the Tour de France, 1950-2008. Note: 1) prizes have been deflated using the Consumer Prize Index published by INSEE. 2) The series are normalized to 1 in 1950.

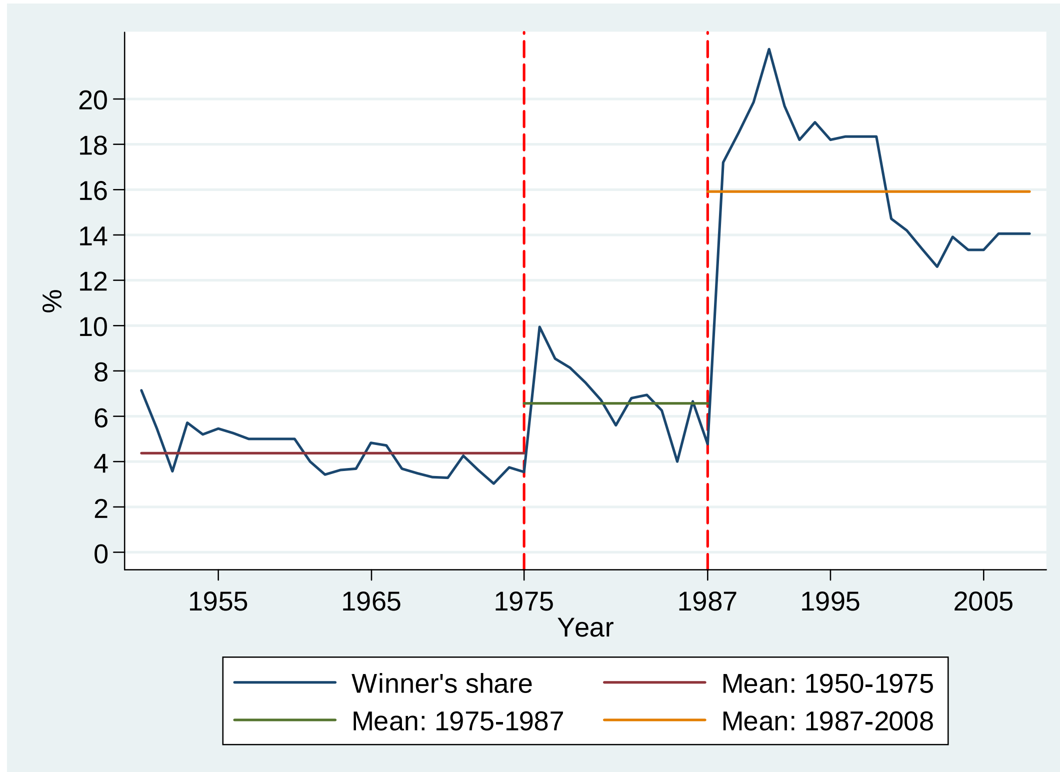

Figure 3: Evolution of the share of total prize money allocated to the Tour de France winner, 1950-2008.

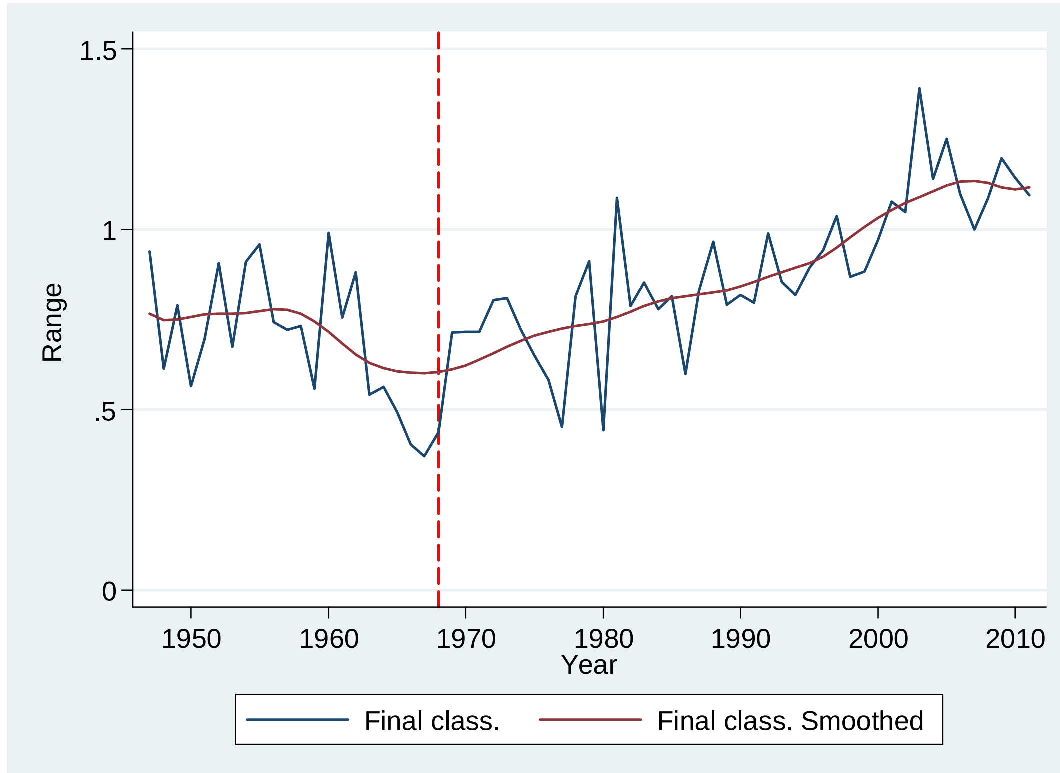

Figure 4: Evolution of the velocity range in the Tour de France, 1947-2011. Note: i) the series are normalized such that 2007=1, ii) the smoothed series are generated using local weighted regression technique on the respective original series with 0.3 bandwidth.

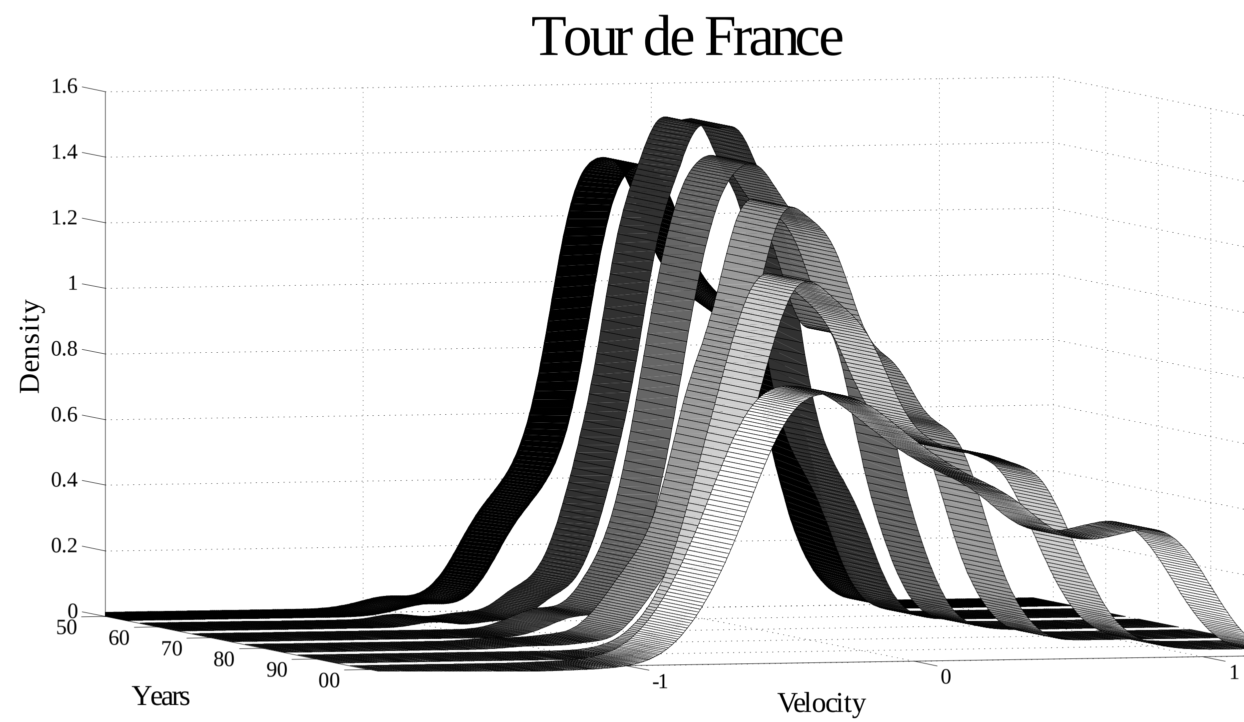

Figure 5: Evolution of the density of velocity in the Tour de France: 1950-2000. Notes: (1) velocity is centered using year specific mean, (2) the density for each decade is generated using kernel methods and pooling data of the 10 corresponding years (8 years for the 2000s).

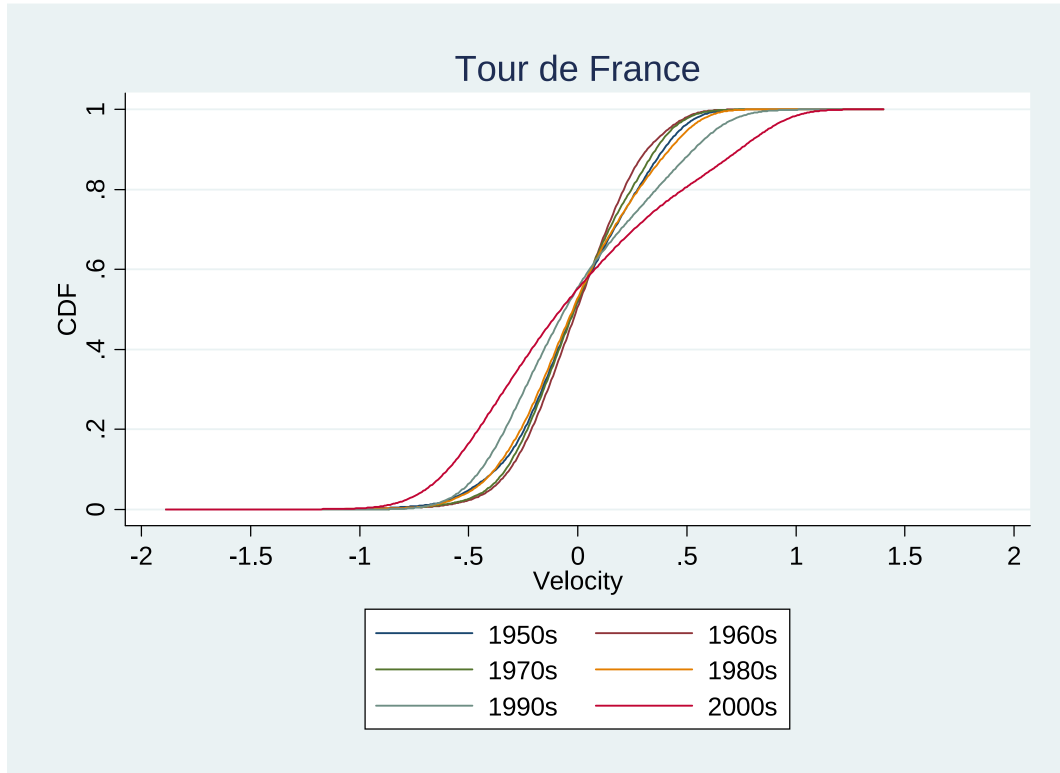

Figure 6: Evolution of the cumulative distribution of velocity in the Tour de France: 1950s-2000s. Notes: (1) velocity is centered using year specific mean, (2) the CDF is based on the density for each decade which is generated using kernel methods and pooling data of the 10 corresponding years (8 years for the 2000s).

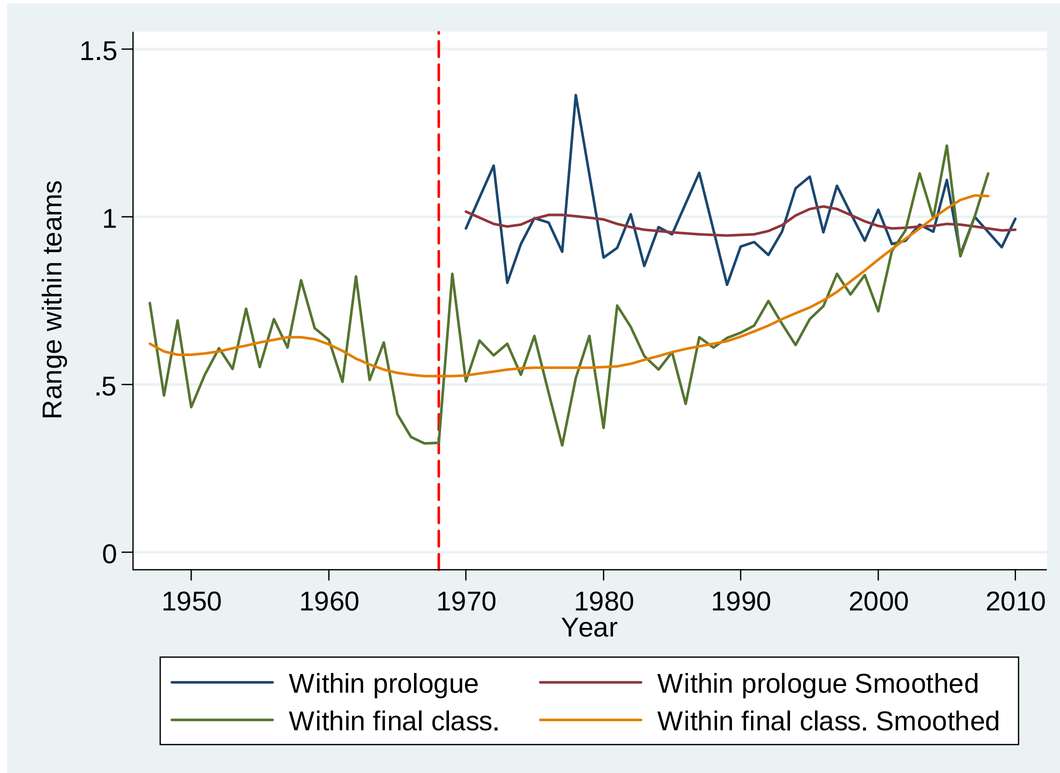

Figure 7: Evolution of the velocity range within team in the final classification and prologue in the Tour de France, 1947-2011. Note: i) the series are normalized such that 2007=1, ii) the smoothed series are generated using local weighted regression technique on the respective original series with 0.3 bandwidth.

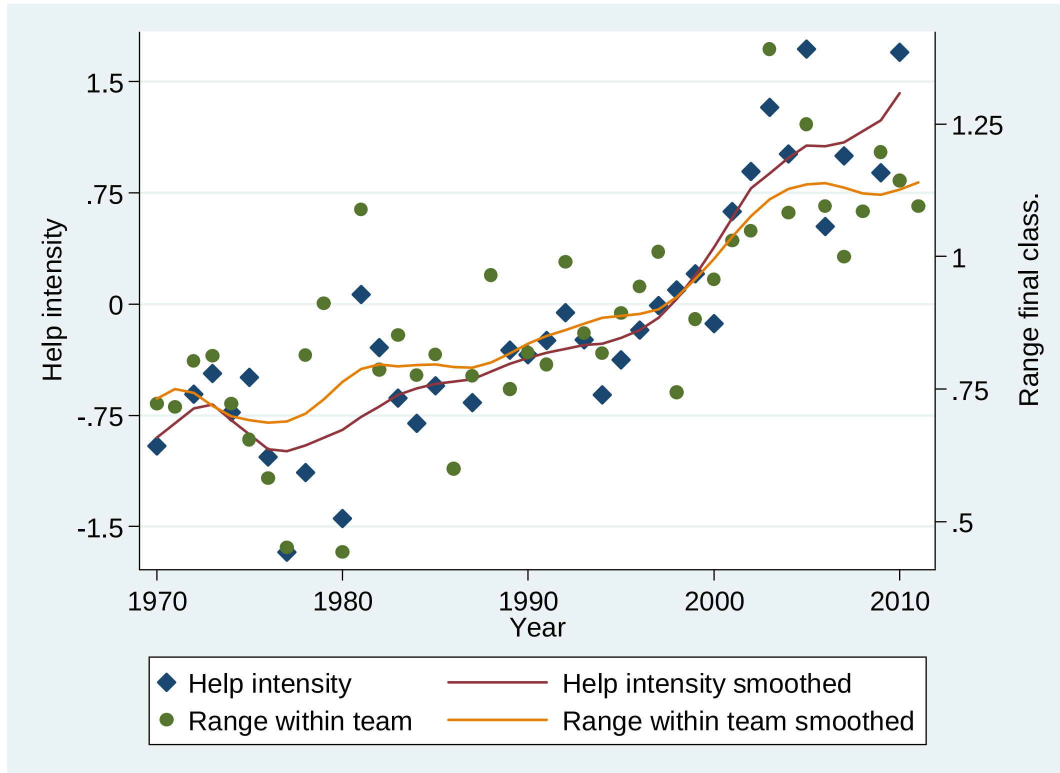

Figure 8: Evolution of the (average) help intensity in the Tour de France, 1970-2011. Note: i) the series are normalized such that 2007=1, ii) the smoothed series are generated using local weighted regression technique on the respective original series with 0.3 bandwidth.