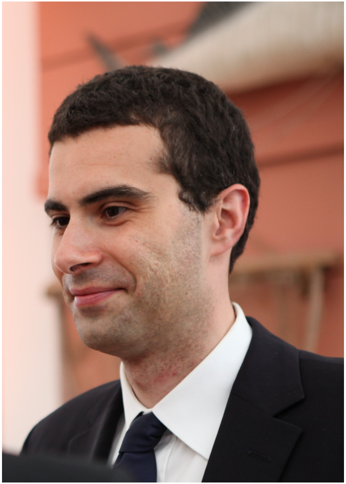
(The unreasonable effectiveness of) Optimal Transport in Economics

Alfred Galichon
New York University and Sciences Po

Work supported by ERC grant CoG-866274 EQUIPRICE

Session 'Frontiers of Modern Econometrics'
World congress of the econometric society
August 17, 2020



In memory of Emmanuel Farhi (1978-2020).

1. Optimal Transport (OT) in a nutshell
2. OT in matching and trade
3. OT in demand estimation and pricing
4. OT in quantile methods

These slides can be downloaded from
alfredgalichon.com/world-congress

Section 1

Optimal transport in a nutshell

- Optimal transport (OT) traces back to Monge in the 18th century and was revived with linear programming (works by Kantorovich, Koopmans, Dantzig and others in the mid 20th century). While the theory mostly arose out of economic motivations (planning problems), it soon drifted away from economics.
- Since the 1990s it has been a very active area again due to connections with physics, dynamic systems and even pure mathematics (Brenier, Villani, McCann).
- Until recently, OT was not part of the standard quantitative economist's toolbox. However, it turns out that many basic problems in economics in a surprisingly diverse fields are *optimal transport problems in disguise*. This connection is useful for theory (existence, uniqueness, stability), and computation (algorithms).
- This talk is a quick overview of some of these connections, and some extensions. It is admittedly skewed toward my own work.

- Optimal transport in the finite case: worker $x \in \mathcal{X}$ matches with firm $y \in \mathcal{Y}$ to produce output $\Phi(x, y)$. Assume $\mathcal{X}$ and $\mathcal{Y}$ are finite and (p_x) and (q_y) are probability vectors over them.
  - Primal problem: form π_{xy} pairs xy in order to maximize total output

$$\max_{\pi \geq 0} \sum_{x,y} \pi_{xy} \Phi_{xy}$$

$$s.t. \begin{cases} \sum_y \pi_{xy} = p_x \\ \sum_x \pi_{xy} = q_y \end{cases}$$

  - Dual problem (by linear programming):

$$\min_{u,v} \sum_{x \in \mathcal{X}} p_x u_x + \sum_{y \in \mathcal{Y}} q_y v_y$$

$$s.t. u_x + v_y \geq \Phi_{xy}$$

- OT is a far-reaching generalization of this duality to the case when $\mathcal{X}$ and $\mathcal{Y}$ are much richer sets. The **Monge-Kantorovich theorem** asserts under general assumptions:
  - Existence of primal solutions (π_{xy})
  - No duality gap: value of primal=value of dual
  - Existence of dual solutions (u_x) and (v_y).

- To facilitate computation, consider the previous primal problem with an entropic regularization in the objective function. Take $\sigma > 0$ a parameter that can be made arbitrarily small. The primal problem

$$\max_{\pi \geq 0} \sum_{x,y} \pi_{xy} \Phi_{xy} - \sigma \sum_{x,y} \pi_{xy} \ln \pi_{xy}$$

$$s.t. \begin{cases} \sum_y \pi_{xy} = p_x \\ \sum_x \pi_{xy} = q_y \end{cases}$$

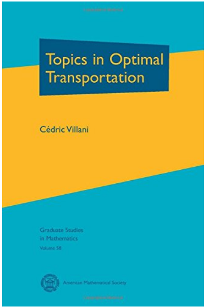
has dual

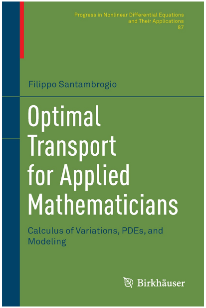
$$\min_{u,v} \left\{ \sum_{x \in \mathcal{X}} p_x u_x + \sum_{y \in \mathcal{Y}} q_y v_y + \sigma \sum_{x,y} \exp\left(\frac{\Phi_{xy} - u_x - v_y}{\sigma} \right) - \sigma \right\}.$$

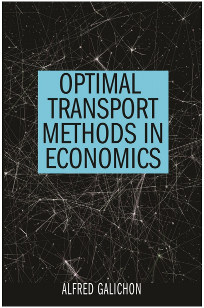
  - Again, extensions beyond finite spaces: *Bernstein-Schrödinger systems*, surveyed in Léonard (2014).
  - Recently, many progress in the computation of this problem: IPFP/Sinkhorn (coordinate descent), mini-batch (stochastic gradient descent), etc.

Monographs

- C. Villani (2004). *Topics in Optimal Transportation*. AMS.
- F. Santambrogio (2015). *Optimal Transport for Applied Mathematicians*. Birkhaüser.
- A. Galichon (2016). *Optimal Transport Methods in Economics*. Princeton.







Section 2

OT in matching and trade

- Consider the decentralized problem of assigning each the distribution of workers (p_x) to the distribution of firms (q_y). A *stable assignment* is a probability mass $(\pi_{xy}) \geq 0$ over $\mathcal{X} \times \mathcal{Y}$ such that there are payoffs vectors u_x and v_y with

$$\begin{cases} \text{counting equations hold: } \sum_y \pi_{xy} = p_x, \sum_x \pi_{xy} = q_y \\ \text{there is no blocking pair: } u_x + v_y \geq \Phi_{xy} \\ \text{feasibility: } \pi_{xy} > 0 \implies u_x + v_y = \Phi_{xy} \end{cases}$$

- **Theorem (Becker-Shapley-Shubik)**. The stable assignments (π_{xy}) are exactly the solutions to the primal OT problem

$$\max_{\pi \geq 0} \sum_{x,y} \pi_{xy} \Phi_{xy}$$

$$s.t. \begin{cases} \sum_y \pi_{xy} = p_x \\ \sum_x \pi_{xy} = q_y \end{cases} . \blacksquare$$

  - This interprets as a welfare theorem: the solution of the central planner coincides with the decentralized equilibrium.
  - Many algorithms to compute this problem efficiently (auction algorithm).

- Consider now π_{xy} =trade flows from country x to country y; p_x=country x's exports; q_y=country y's imports
  Theorem (Alan Wilson). The gravity equation in trade

$$p_x = \sum_y \exp\left(\Phi_{xy} - u_x - v_y\right)$$
$$q_y = \sum_x \exp\left(\Phi_{xy} - u_x - v_y\right)$$

  (where Φ_{xy} gravity term, and u_x and v_y= exporter and importer fixed effects) is the solution to the dual regularized optimal transport problem

$$\min_{u,v} \left\{ \sum_{x \in \mathcal{X}} p_x u_x + \sum_{y \in \mathcal{Y}} q_y v_y + \sum_{x,y} \exp\left(\Phi_{xy} - u_x - v_y\right) \right\}. \blacksquare$$

  - **Intuition**: first order conditions.
  - Fast **computation** via coordinate descent.

- Estimation problem: recover "matching surplus" / "propensity to trade" Φ_{xy} based on observation of trade flows $\hat{\pi}_{xy}$. Parameterize $\Phi^{\lambda}_{xy} = \sum_k \lambda_k \phi^k_{xy}$.
  Theorem (Dupuy-Galichon-Sun). The unique λ satisfying

$$\sum_{y\in\mathcal{Y}} \pi^{\lambda}_{xy} = \sum_{y\in\mathcal{Y}} \hat{\pi}_{xy} =: p_x, \quad \sum_{x\in\mathcal{X}} \pi^{\lambda}_{xy} = \sum_{x\in\mathcal{X}} \hat{\pi}_{xy} =: q_y, \sum_{x,y} \pi^{\lambda}_{xy}\phi^k_{xy} = \sum_{x,y} \hat{\pi}_{xy}\phi^k_{xy}$$

  is the solution to

$$\min_{u,v,\lambda} \left\{ \sum_{x\in\mathcal{X}} p_x u_x + \sum_{y\in\mathcal{Y}} q_y v_y + \sum_{x,y} \exp\left(\Phi^{\lambda}_{xy} - u_x - v_y\right) - \sum_{xy} \hat{\pi}_{xy}\Phi^{\lambda}_{xy} \right\} . \blacksquare \tag{1}$$

  - **Intuition:** Optimality wrt λ: matching moments of ϕ^k; optimality wrt u_x and v_y: matching the right marginals.
  - **Remark**: Problem (1) is a pseudo-poisson MLE (PPML) with x and y fixed effects introduced in trade by Santos Silva and Tenreyro (2006).
  - **Estimation of matching models**: Choo and Siow (2006), Galichon and Salanié (2020), Dupuy and Galichon (2014).

- There are T periods. If xy are matched at period t, then x's type transitions to x' with probability $P_{x'|xy}$ and y's transitions to y' with probability $Q_{y'|xy}$. Period 1 distributions of types are (p_x) and (q_y). **Theorem (Fox-Galichon-Corblet)**. The central planner 's problem

$$\max_{\pi_{xy}^t \geq 0} \sum_{xy} \pi_{xy}^t \Phi_{xy}^t$$

$$s.t. \begin{cases} \sum_y \pi_{xy}^1 = p_x, \ \sum_x \pi_{xy}^1 = q_t \\ \sum_{y'} \pi_{x'y'}^t = \sum_{xy} P_{x'|xy} \pi_{x'|xy'}^{t-1}, \ t \geq 2 \\ \sum_{x'} \pi_{x'y'}^t = \sum_{xy} Q_{y'|xy} \pi_{x'|xy'}^{t-}, \ t \geq 2 \end{cases}$$

can be decentralized via the dual problem

$$\min \sum_x p_x U_x^1 + \sum_y q_y V_y^1$$

$$s.t. \begin{cases} U_x^t + V_y^t \geq \Phi_{xy}^t + \sum_{x'} U_{x'}^{t+1} P_{x'|xy} + \sum_{y'} V_{y'}^{t+1} Q_{y'|xy}, \ t \leq T-1 \\ U_x^T + V_y^T \geq \Phi_{xy}^T \end{cases} \quad .\blacksquare$$

  - **Structural estimation** of this model in Ciscato, Dupuy, Fox, Galichon and Weber (2020).

Section 3

OT in demand estimation and pricing

- Consider the (additive) discrete choice problem

$$u_i = \max_j \left\{ V_j + \varepsilon_{ij} \right\}, i \in \{1, ..., n\}$$

the problem of *discrete choice inversion* consists of determining the systematic utility V_j based on the market share $q_j =$ frequency of j chosen.

Theorem (Galichon-Salanié). The solutions (u, v) to the dual OT problem

$$\min_{u,v} \sum \frac{1}{n} u_i + \sum q_j v_j$$
$$s.t.\ u_i + v_j \geq \varepsilon_{ij}$$

solve the discrete choice inversion problem above with $V_j = -v_j$.■

  - **Intuition**: by duality, "consumers choose yogurts" $<=>$ "yogurts choose consumers" $<=>$ "consumers match with yogurts".
  - **Consequence:** use of OT algorithms for discrete choice inversion. Nonsmooth models: Chiong, Galichon and Shum (2017); continuous models: Chernozhukov, Galichon, Henry and Pass (2020).

- Consider now mixed logit model (Berry, Levinsohn, Pakes)

$$u_i = \max_j \{ V_j + \varepsilon_{ij} + \sigma\eta_j \}, i \in \{1, ..., n\}$$

where η_j is EV-type I, and ε_{ij} is e.g. pure characteristics $\varepsilon_{ij} = \epsilon_i^\top \xi_j$.
Theorem (Bonnet-Galichon-Hsieh-O'Hara-Shum). The solutions (u, v) to the dual regularized optimal transport problem

$$\min_{u,v} \sum_i \frac{1}{n} u_i + \sum_j q_j v_j + \sigma \sum_{ij} \exp\left(\frac{\varepsilon_{ij} - u_i - v_j}{\sigma}\right)$$

identify with the solutions to the discrete choice inversion problem with $V_j = -v_j$.$\blacksquare$

  - The coordinate descent algorithm coincides BLP's contraction mapping algorithm.
  - Advantage of the reformulation: (1) compute mixed logit ($\sigma > 0$) and pure characteristics ($\sigma = 0$) all at once; and (2) extend the approach to nonadditive case.

- Consider a quasilinear hedonic model where each producer $x \in \mathcal{X}$ produces one unit of good and chooses in which quality $z \in \mathcal{Z}$. Each consumer $y \in \mathcal{Y}$ consumes one unit of good, and chooses in which quality $z \in \mathcal{Z}$. The distribution of the producers and consumers are (p_x) and (q_y).
- Hedonic equilibrium (Ekeland-Heckman-Nesheim, 2004): for each $z \in \mathcal{Z}$, supply for z = demand for z. Letting P_z be the price of quality z, the producer's and consumer's problem are respectively

$$u_x = \max_{z \in \mathcal{Z}} \{P_z - C_{xz}\} \text{ and } v_y = \max_{z \in \mathcal{Z}} \{U_{yz} - P_z\}$$

Theorem (Chiappori-McCann-Nesheim). Setting $\Phi_{xy} = \max_z \{U_{yz} - C_{xz}\}$, the indirect utilities u_x and v_y at equilibrium are the solutions to optimal tranport problem

$$\max_{\pi \geq 0} \sum_{xy} \pi_{xy} \Phi_{xy}$$

$$s.t. \begin{cases} \sum_y \pi_{xy} = p_x \\ \sum_x \pi_{xy} = q_y \end{cases}$$

and the equilibrium prices P_z can be deduced from u_x and v_y.■

Section 4

OT in quantiles methods

- Let (X, Y) be a random vector over $\mathbb{R}^{d+1}$. The conditional quantile $Q_{Y|X}(t|x)$ of Y given $X = x$ is the inverse of the c.d.f. $F_{Y|X=x}(.|x)$, that is $q = Q_{Y|X}(t|x)$ if and only if

$$\Pr(Y \leq q | X = x) = t.$$

Theorem (Carlier-Chernozhukov-Galichon). One has

$$Q_{Y|X}(t|x) = \frac{\partial \varphi}{\partial t}(x, t)$$

where φ is the solution to the optimal transport problem

$$\max_{\varphi,\psi} \int \varphi(x, t)\, dt dF_X(y) + \int \psi(x, y)\, dF_{XY}(x, y)$$
$$s.t.\ \varphi(t, x) + \psi(x, y) \geq ty. \blacksquare$$

  - Application to **multivariate quantiles**: when Y is multivariate, replace the product by the scalar product to get a notion of multivariate quantiles. This found applications to risk measures by Ekeland, Galichon and Henry (2012), decision theory in Galichon and Henry (2012), multivariate depth in Hallin, Chernozhukov, Galichon and Henry (2017).

- Assume a parametric form of the conditional quantile

$$Q_{Y|X}(t|x) = x^\top \beta(t). \qquad (2)$$

  Quantile regression (Koenker and Bassett 1978, Koenker 2005) is about the estimation of $\beta(t)$.
- **Theorem (Carlier-Chernozhukov-Galichon)**. If (2) holds, then

$$\beta(t) = b'(t),$$

  where b is obtained from the solving the following extension of the optimal transport problem

$$\max_{\varphi,\psi} \int x^\top b(t)\, dt dF_X(y) + \int \psi(x,y)\, dF_{XY}(x,y) \qquad (3)$$
$$s.t.\ x^\top b(t) + \psi(x,y) \geq ty. \blacksquare$$

  - In the scalar case (3) can be interpreted as a shaped-constrained quantile regression. In the multivariate case, allows to get a multivariate extension of quantile regression, see Carlier, Chernozhukov, Galichon (2016, 2017), Carlier, Chernozhukov, De Bie, Galichon (2020).

Section 5

Perspectives

- We saw several problems arising in economics that had the structure of optimal transport. Useful for:
  - theory: use of optimal transport and convex analysis for existence, uniqueness, stability in these problems
  - computation: use of numerical optimal transport
- We reviewed two extensions of optimal transport:
  - Dynamic optimal transport => dynamic matching with transferable utility
  - Vector quantile regression => parametric estimation of multivariate quantile

- Despite its wide applicability, optimal transport is intrinsiquely connected with transferable utility / quasilinear utility / optimization, and fails to cover a number of important problems:
    - Matching without transfers (Gale-Shapley)
    - Matching with imperfectly/partially transferable utility (taxes, salary caps, etc.)
    - Nonadditive random utility models
    - Hedonic models beyond quasi-linear utility
    - Dynamic programming beyond linear programming
- To handle these topics, a more general framework is needed, the *equilibrium flow problem.* Embeds optimal transport and the above topics.
    - Current work with Larry Samuelson and Lucas Vernet.

Section 6

Reference

Bonnet, Galichon, O'Hara, Hsieh and Shum (2020). Yogurts choose consumers? Identification of Random Utility Models via Two-Sided Matching. *SSRN*.

Carlier, Chernozhukov, De Bie, Galichon (2020). Vector quantile regression and optimal transport, from theory to numerics. *Empirical Economics*.

Carlier, Chernozhukov, Galichon (2016). Vector quantile regression: an optimal transport approach. *Annals of statistics*.

Carlier, Chernozhukov, Galichon (2017). Vector quantile regression beyond correct specification. *Journal of Multivariate Analysis*.

Chernozhukov, Galichon, Henry and Pass (2020). Single market nonparametric identification of multi-attribute hedonic equilibrium models. *Journal of Political Economy* (forthcoming).

Chiappori, McCann, Nesheim (2009). Hedonic price equilibria, stable matching, and optimal transport: equivalence, topology, and uniqueness. *Economic Theory*.

Chiong, Galichon, Shum (2016). Duality in dynamic discrete choice models. *Quantitative Economics*.

Choo, Siow (2006). Who marries whom and why. *Journal of Political Economy*.

Ciscato, Dupuy, Fox, Galichon, Weber (2020). Family dynamics: marriage, fertility and divorce. In progress.

Dupuy, Galichon (2014). Personality traits and the marriage market. *Journal of Political Economy*.

Dupuy, Galichon, Sun (2019). Estimating matching affinity matrix under low-rank constraints. *Information and Inference*.

Ekeland, Heckman, Nesheim, 2004. Identification and estimation of hedonic models. *Journal of Political Economy*.

Ekeland, Galichon, Henry (2012) Comonotonic measures of multivariate risks. *Mathematical Finance*.

Fox, Galichon, Corblet (2020). A dynamic model of two-sided matching. In progress.

Galichon (2016). *Optimal Transport Methods in Economics*. Princeton.

Galichon, Henry (2012). Dual theory of choice with multivariate risks. *Journal of Economic Theory*.

Galichon, Salanié (2020). Cupid's invisible hands: social surplus and identification in matching models. *SSRN*.

Galichon, Samuelson, Vernet (2020). The Equilibrium Flow Problem. In progress.

Reference for this presentation (3/3)

Hallin, Chernozhukov, Galichon and Henry (2017). Monge-Kantorovich Depth, Quantiles, Ranks and Signs. *Annals of Statistics*.

Koenker (2005). *Quantile regression*. Cambridge.

Koenker, Bassett (1978). Regression Quantiles. *Econometrica*.

Léonard (2014). A survey of the Schrödinger problem and some of its connections with optimal transport. *Discrete & Continuous Dynamical Systems*.

Santambrogio (2015). *Optimal Transport for Applied Mathematicians*. Birkhaüser.

Santos Silva and Tenreyro (2006). The log of gravity. *Review of Economics and Statistics*.

Villani (2004). *Topics in Optimal Transportation*. AMS.