“A game-theoretic analysis of childhood vaccination behavior: Nash versus Kant”*

by

Philippe De Donder, Toulouse School of Economics (CNRS)†

Humberto Llavador, Universitat Pompeu Fabra and Barcelona School of Economics‡§

Stefan Penczynski, University of East Anglia**

John E. Roemer, Yale University††

Roberto Vélez, Centro Estudios Espinosa Yglesias, Mexico City‡‡

September 2022

Supporting Information

Additional Supporting Information can be found in the Online Appendix at https://bit.ly/3xrHJ3R

* We thank Peter Aronow, Nicholas Christakis, Alan Gerber, Mikhael Guy, Markus Jantti, Walter Orenstein, Nicolas Pistolesi, François Poinas, Kenneth Scheve, Joaquim Silvestre, Milan Svolik, Roberto Veneziani, Rodolfo de la Torre and seminar participants at the Aix-Marseille School of Economics, Lille University, Paris School of Economics and Brown University for their advice and comments. Collin Schumock, Austin Jang, and Joon Lee, all from Yale University, provided excellent research assistantship.

† Philippe De Donder acknowledges the French ANR under grant ANR-17-EURE-0010 (Investissements d'Avenir program) for financial support. philippe.dedonder@tse-fr.eu

‡ Humberto Llavador acknowledges financial support from the Spanish Agencia Estatal de Investigación (AEI), through the Severo Ochoa Programme for Centres of Excellence in R&D (Barcelona School of Economics CEX2019-000915-S) and grants AEI/FEDER,UE-PGC2018-096370-B-I00 and PID2020-115044GB-I00//AEI/10.13039/501100011033. humberto.llavador@upf.edu

§ Corresponding author

** s.penczynski@uea.ac.uk

†† John Roemer acknowledges the Institute for Social and Policy Studies at Yale University for financial support. john.roemer@yale.edu

‡‡ rvelezg@ceey.org.mx

Abstract

The vaccination game exhibits positive externalities that may lead to herd immunity. The standard game-theoretic approach assumes that parents make decisions according to the Nash behavioral protocol, which is individualistic and non-cooperative. However, in more solidaristic societies, parents may behave cooperatively, optimizing according to the Kantian protocol, in which the equilibrium is efficient. We test whether childhood vaccination behavior conforms better to the individualistic or cooperative protocol. The Kant model dominates the Nash model on our six-country sample. We conjecture that a social norm evolved, and that the Kantian equilibrium offers one precise version of such a social norm.

Keywords: Kantian equilibrium, Nash equilibrium, vaccination, social norm.

JEL code: C72 D62 D63 I12

1 Introduction

Vaccination against childhood diseases has improved child health and life expectancy dramatically over the last fifty years. Researchers from the US Center for Disease Control and Prevention (CDC) and the World Health Organization (WHO) write that in 2017, 110,000 children died of measles infection globally, and that in the period 2000-2017, 21 million lives were saved by measles vaccination (Dabbagh, Laws et alii, 2018, Table 2). The fraction of children globally who are vaccinated against measles rose in this period from 72% to 85%. Sweden and China have vaccination coverage rates in 2019 of 97% and 99%, respectively (World Bank, https://data.worldbank.org/indicator/SH.IMM.MEAS?view=map).

Our interest in this article, however, is not epidemiological, but rather behavioral. Vaccination is a choice in which cooperation among the population is important. A child's vaccination provides a positive externality for others, because as the vaccination coverage rate increases, the probability that an unvaccinated child contracts the disease decreases. Eventually 'herd immunity' may be attained, when the coverage rate is sufficiently high that the virus can find no hosts in the population.

Here, we model vaccination behavior as a game, in which the strategy of parents is to choose whether or not to vaccinate their child, or, in a more general version, a parent's mixed strategy is a probability that she will vaccinate her child.

We summarize the central concepts of this paper, Nash and Kantian equilibrium.

Definition 1 Let $\boldsymbol{V} = \{V_1, V_2 \ldots, V_n\}$ be a set of payoff functions for *n* players, where the strategy space for each player is the unit interval *I* and for all *j*, $V_j: I^n \rightarrow \mathbb{R}$. An *n*-tuple of strategies $a = (a_1, a_2 \ldots, a_n)$ where a_i is the probability that parent *i* vaccinates her child, is a *Nash equilibrium* of the game if, for all $j = 1, \ldots, n$:

$$a_j \in arg\max_{x \epsilon I} V_j(a_1, \ldots, a_{j-1}, x, a_{j+1,\ldots,} a_n).$$

Define for any number $x \in I$ and any number $\rho \geq 0$ the truncation:

$$\rho \circ x = \min[\rho x, 1].$$

If a is a strategy profile, denote the mean of the function $\rho \circ a$ by $\overline{\rho \circ a}$.

A *multiplicative Kantian equilibrium* is a profile of strategies $a = (a_1, \dots, a_n) \epsilon I^n$ such that no player would prefer, for some non-negative factor ρ, the truncated rescaled profile $\rho \circ a \equiv (\rho \circ a_1, \dots, \rho \circ a_n)$:

$$\text{For all } j,\ 1 = arg \max_{0 \le \rho \le 1/a_j} V_j(\rho \circ a).$$

The truncated re-scaled profile is a vector of probabilities.

A picture provides some intuition. In Figure 1, we depict a possible Nash or Kantian equilibrium L=(L_1, L_2) in a game with two players. Suppose the strategy space for each player is the non-negative real line. In Nash optimization the column player 2 examines the set of counterfactual profiles consisting of the dashed vertical line through *L*, and the row player 1 examines the counterfactual profile of strategies where only he deviates, which is the horizontal dotted line through *L*. In contrast, the Kantian players – both row and column – examine the *same* set of counterfactual profiles to test for an equilibrium, which is the ray through *L*. The mathematical expression of cooperation captured by Kantian optimization is that the players always examine a common set of counterfactual profiles. If you will, the players are acting in concert. In contrast each player in Nash optimization is 'going it alone—' he treats the other player(s) as part of his environment, not as part of the action.

Definition 2 A game, as defined in definition 1, is *strictly monotone increasing (decreasing)* if each player's payoff function is strictly increasing (decreasing) in the strategies of the other players.

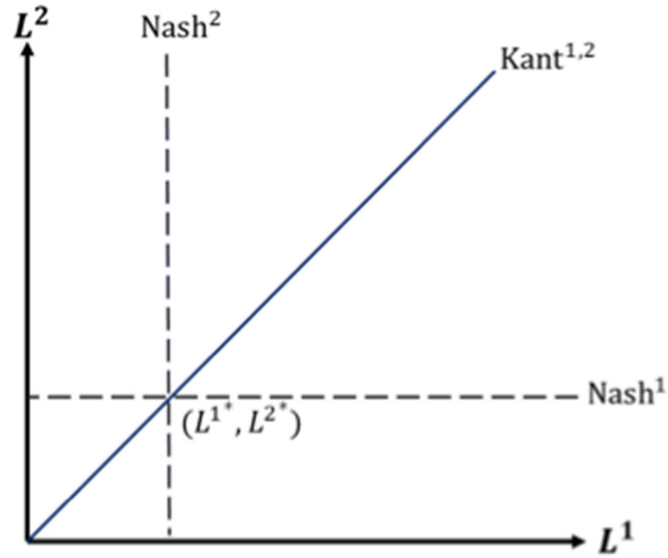

Figure 1. The set of counterfactuals in a Nash and a Kantian equilibrium. The picture shows that, unlike Nash players, Kantian players share the set of deviations they contemplate.

We have:

<u>Proposition 1</u> *Any interior Nash equilibrium of a strictly monotone game where all payoff functions are differentiable is Pareto inefficient.*

<u>Proposition 2</u> *Any strictly positive multiplicative Kantian equilibrium of a strictly monotone game is Pareto efficient.*

The proof of Proposition 1 is provided in Appendix A. Proposition 2 is proposition 3.1 in (Roemer, 2019, p. 42). The Pareto inefficiency of Nash equilibrium in monotone increasing games is called, in the vernacular, the *free-rider problem*, whereas its inefficiency in monotone decreasing games is called the *tragedy of the commons.* Thus, the content of Propositions 1 and 2 is that cooperation, conceived of as Kantian optimization, resolves the free-rider problem and the tragedy of the commons which are ubiquitous in Nash equilibrium.

Note that the Kantian optimization protocol does not rely on the altruism of parents. When a parent examines re-scalings of the proposed profile, she is forced to take into account the external effect on her own welfare brought about by the actions of others. In this way, she internalizes the externality. Think of the question a citizen asks herself when contemplating whether to make a small increase in her contribution to construction or financing of a public good. A Nash player asks herself whether her own disutility from increasing her contribution is worth the small increment in the size of the public good to her. She may well decide not to contribute under the Nash protocol. But a Kantian player asks, "How would I like it if *everyone* increased his contribution to the public good in like manner?" She tests the positive externality by asking what effect her increased contribution, if emulated by everyone, would have upon her welfare. These two different approaches are represented in Figure *1*. The question the Kantian player poses induces her to take into account the positive externality of vaccination. The consequence, though perhaps not obvious, is that Pareto efficiency is achieved in the Kantian equilibrium.[1]

In the Kantian approach, we alter the way that players optimize in a game, but retain classical self-interested preferences. In contrast, much of behavioral economics alters preferences, but retains Nash optimization. In the latter, arguments like the welfare of others, fairness, warm glows, etc. are added to the domain of preferences. In the former, a cooperative or fairness ethic is embodied in the manner of optimizing, not in preferences. Both approaches find that the level of vaccination in fact exceeds what is predicted by purely self-interested behavior. One virtue of the Kantian approach is that it embeds a precisely defined social norm thus enabling us to run a horse race of Nash against Kant, because the two concepts use the same data.

We can now state our project in this article. We will calibrate the parameters of the vaccination game using surveys that interview parents in a set of countries about their beliefs about the costs and benefits of vaccination. We will then compute the Nash equilibrium and the multiplicative Kantian equilibrium of the game in each country. This will consist of two profiles of vaccination probabilities in the country, and their implied equilibrium coverage rates. We will then ask which of these equilibria appears to better explain observed vaccination behavior in the country. Do parents appear to be 'going it alone' or cooperating? Of course, the reality is surely that some people go it alone and some behave

[1] For a study of Kantian equilibrium, see Roemer (2019).

cooperatively, but we will not attempt to analyze a model that is so nuanced: we will be satisfied with the simpler question just posed.

It must be emphasized that we are not attempting to achieve the best possible fit of a fully articulated model of vaccination to the data. Such a model would need many more parameters than ours has. Our goal is to look for the evidence concerning whether childhood vaccination behavior of parents is better described as cooperative or going it alone. Our interest is not in vaccination as such, but rather in whether people participating in a project with externalities appear to treat them in a cooperative manner, in order to prevent free-rider problems and commons' tragedies.

We now briefly review related works. One of the first papers credited in the economic literature with studying the insufficient immunization rates due to the incomplete internalization of the positive vaccination externality is Brito *et al.* (1991). Geoffard and Philipson (1997) are the first to study the forces that make disease eradication through vaccination difficult in the context of a dynamic, SIR, model (where individuals are either susceptible, infected or immune through recovery). The SIR model has since proved quite popular in the economic literature (see for instance Auld (2003) and Philipson (2000) for an early survey of this literature).

Several recent papers in the economics literature rather model the individual choice to vaccinate in a static setting with one or two periods. They use a decision theoretic approach where individuals choose whether to vaccinate as a function of the disease prevalence (or fraction of the population vaccinated), without strategic interactions between agents. They differ in whether vaccines are perfect (i.e., prevent the occurrence of the disease for sure) or not, in the vaccination costs (financial costs, time costs and/or side effects) and in whether prevention efforts (such as masks for instance) are available or not. For instance, Nuscheler and Roeder (2016) study the impact of time preferences on the choice to vaccinate, while Crainich et al (2019) concentrate on risk aversion. d'Albis et al (2022) study the impact of pessimistic expectations on vaccination decisions.

The use of a game theoretical approach to the vaccination decision is more common in the epidemiology literature. The first study of vaccination behavior with a game theoretical perspective was prompted by concerns associated with the pertussis vaccine (Fine et al, 1976). Since then, epidemiological game theory models have been formulated for several diseases, including measles (Shim et al, 2012b). Papers in this literature merge together a population-level epidemiological model for the disease transmission (à la SIR) and an individual-level calculation of payoff associated with infection and/or vaccination. These studies repeatedly show that the pursuit of self-interest would lead to suboptimal vaccination coverage for a community. (See Bauch et al (2003, 2004) for accessible examples of this literature).

All papers above assume that agents are self-interested. The paper closest to ours is Shim et al. (2012a), who first build a simple game theoretical model where agents may exhibit some altruism. By varying the degree of altruism, one moves from the selfish Nash equilibrium with too little vaccination to the socially optimal behavior. They then resort to a survey to elicit the beliefs of agents regarding the parameters of their model, such as the efficacy and dangerosity of the influenza vaccines, as well as their perceived risk of infection (risk-to-self) and of transmitting the disease (risk-to-others). They then estimate an econometric model of the individuals' decisions to vaccinate and compute the agent's degree of altruism as the ratio of the coefficients of the risk-to-other divided by the risk-to-self. They obtain a baseline value of the degree of altruism of 0.25. They then compute and compare the vaccination rates for perceived parameters at the selfish equilibrium (27%), with the baseline degree of altruism (34%) and at the social optimal (46%). They also find that, for any altruism degree, the vaccination coverage is lower with the true parameters than with perceived parameters, presumably because people tend to overestimate their infectious period as well as their infection probability. So, the lack of altruism leads to too little vaccination, while perception errors lead to too much vaccination, but with the former effect being much larger than the latter.

In health behavior studies, the relevance of social norms and cooperative attitudes is acknowledged in general (Vanlandingham et al. 1995) and specifically for the case of vaccination (Yang 2015). As they are rather homogeneous and not easily changed, social norms are not a prime factor in studies that aim to locally predict or influence behavior (e.g. Kreps et al. 2020, Gatwood et al. 2021). Here, more prominent variables are vaccine-related attributes such as side effects, vaccine safety and efficacy as well as political factors such as health authority approval, endorsements, party affiliation and origin of vaccine. However, in order to fully and globally understand vaccination behavior, social norms are considered important (Kan and Zhang 2018).

Section 2 presents a random utility model of vaccination and the equilibrium theory. Section 3 describes the data. Section 4 presents our method of estimation, that is, of testing whether the Nash or Kantian model better explains vaccination behavior in six countries. Section 5 presents our major finding: in all six countries, the Kantian model performs significantly better than the Nash model in explaining behavior. Section 6 argues that the reason this is true is that, due to the existence of positive externalities and the free rider problem, a social norm has evolved to vaccinate one's children, and this engenders vaccination rates that are uniformly greater than those predicted by Nash equilibrium. Section 7 offers a short conclusion. The Online Appendix presents details that are elided in the main text.

2 A random utility model of vaccination behavior

2.1 The set-up

We model the problem of the parent who must decide whether or not to vaccinate her child against measles. We assume there are no laws or regulations mandating vaccination.[2] If the child is not vaccinated, there is the possibility that he will contract measles and possibly die or suffer a debilitating illness. If he is vaccinated, he will either be healthy and protected from measles, or may suffer a side effect from the vaccination of some severity (or so the parent may believe).

We define three states of the child's health: healthy (H), suffering a possibly severe side effect from an inoculation (E), or contracting measles and possibly suffering a very severe outcome or death (D). Table

[2] In all six countries, at the time vaccination occurred, there was no legal requirement to vaccinate one's child. See p. 19 below.

1 presents the parent's beliefs about probabilities of the three health states if vaccinated and if not vaccinated, and the von Neumann –Morgenstern utilities of the parent (the decision maker) based upon the child's health outcome.

Table 1. Utilities and probabilities of health states. Columns represent the three possible states of a child's health. Rows show the utility, and the probabilities of each state for a child who was not and who was vaccinated.

	Healthy	Side effect	Death/Severe Disability
Utility	1	u	0
Probability if not vax	$1-p_0$	0	p_0
Probability if vax	$1-p$	p	0

The utilities of the states Healthy and Death are a normalization that fixes the von Neumann-Morgenstern utility function of the parent. The utility u from the possible side effect is strictly between zero and one. We call the ordered pair (p,u) the parent's *type*; it is her beliefs about the utility-relevant facts concerning the side effect of vaccination. p_0 is the probability of death or severe disability conditional upon contracting measles. We will take p_0 to be common knowledge of parents. We assume the population is characterized by a Beta distribution Q of (p,u) defined on the unit square. (Thus, we assume that $0<u<1$ for all types.) Parent i's *mixed strategy* will be a number $a_i \in [0,1]$, the probability with which she will vaccinate her child. The parent's (von Neumann Morgenstern) expected utility is defined on the ordered pair $(a_i, \bar{a})$, where $\bar{a}$ is the *coverage rate* in the population, defined as the average probability of vaccination across all parents. There is a *probability function* $\pi: [0,1] \to [0,1]: \bar{a} \to \pi(\bar{a})$ which gives the probability that an unvaccinated child will contract measles if the coverage rate is $\bar{a}$. The positive externality is modeled by supposing that π is a strictly decreasing, continuous function. Expected utility for a parent of type (p,u) is given by:

$$V_{(p,u)}(a,\bar{a}) = a\overbrace{\big((1-p)\cdot 1 + pu + \varepsilon\big)}^{ex\ utility\ if\ vacc} + (1-a)\overbrace{\Big(\pi(\bar{a})(1-p_0) + \big(1-\pi(\bar{a})\big)\Big)}^{ex\ utility\ if\ not\ vacc} \qquad (2.1)$$

The number ε is the realization of a random variable with a distribution function $L(\cdot)$ on support $\mathbb{R}$, which is drawn i.i.d. across all parents. It is assumed that 99% of the support of L lies on the non-negative real numbers: this models a positive utility saltus that the parent receives if she vaccinates her child – because she is doing what physicians and society recommend, what most of her neighbors are doing, and so on. The main motivation for inserting this random element into utility is that it will guarantee that the Nash and Kant equilibrium strategies all lie in the open interval (0,1), a property that is essential for our estimation strategy (see Section 4 below). Thus, the data of the problem are $\{Q(\cdot), p_0, \pi(\cdot), L(\cdot)\}$.

We call the set of parents of a given type (p, u) a *tranche.* It is assumed, in particular, that ε is distributed i.i.d. within every tranche. This means that we will observe the behavior of a type (p, u) as a mixed strategy, even if every member of the tranche has a pure strategy, as long as the effect of the random variate differs across individuals. The statistician will only observe the average probability of vaccination within each tranche, which we will denote $a(p, u)$. This is to be thought of as the fraction of those of type (p, u) who decide to vaccinate, depending on their draw of the utility bump ε.

2.2 Nash equilibrium

A *Nash equilibrium of the game*, given a realization of the random variate *L*, is an action of 'vaccinate' or 'do not vaccinate' for every individual within every type given by:

$$vaccinate = \begin{cases} 1, \ if \ \frac{dV_{(p,u)}}{da} \equiv p(u-1) + p_0\pi(\bar{a}^N) + \varepsilon > 0 \\ 0, \ if \ \frac{dV_{(p,u)}}{da} \equiv p(u-1) + p_0\pi(\bar{a}^N) + \varepsilon < 0 \end{cases}, \tag{2.2}$$

where $\bar{a}^N$ is the fraction of individuals who vaccinate in equilibrium[3]. Formally, we say that the Nash equilibrium is a strategy $\alpha^N(p, u, \varepsilon)$ for each individual (p, u, ε) and a coverage rate $\bar{a}^N$ such that:

$$\alpha^N(p, u, \varepsilon) = \begin{cases} 1 \quad if \quad \varepsilon > p(1-u) - p_0\pi(\bar{a}^N) \\ 0 \quad if \quad \varepsilon < p(1-u) - p_0\pi(\bar{a}^N) \end{cases} \tag{2.3}$$

and:

[3] We can ignore the null set of types for which $p(u-1) + p_0\pi(\bar{a}^N) + \varepsilon = 0$

$$\bar{a}^N = \int_{(p,u)} \int_{p(1-u)-p_0\pi(\bar{a}^N)}^{\infty} dL(\varepsilon)dQ(p,u) = \int \Big(1 - L\big(p(1-u) - p_0\pi(\bar{a}^N)\big)\Big)\, dQ(p,u) \quad (2.4)$$

□ Thus, vaccinate if and only if $\varepsilon > p(1-u) - p_0\pi(\bar{a}^N)$, an event that occurs (in the (p,u) tranche) with probability $1 - L\big(p(1-u) - p_0\pi(\bar{a}^N)\big)$. The fraction of this tranche that vaccinates is:

$$a^N(p,u) = \int_{p(1-u)-p_0\pi(\bar{a}^N)}^{\infty} \alpha^N(p,u,\varepsilon)dL(\varepsilon) = 1 - L\big(p(1-u) - p_0\pi(\bar{a}^N)\big) \quad (2.5)$$

Equation (2.4) is a single equation in the unknown $\bar{a}^N$. We solve it for $\bar{a}^N$, and then compute the Nash equilibrium strategy profile from equation (2.5). Note that it appears as if the type (p,u) has a (single) mixed strategy, $a^N(p,u)$.

It is also noteworthy that, because the support of L is the entire real line, every Nash strategy is in the open interval $(0,1)$. See equation (2.5). This will be an important fact in what follows.

2.3 Kantian equilibrium

It will be convenient to define:

$$g(u,\bar{a}) = \frac{p_0\pi(\bar{a})}{1-u}. \quad (2.6)$$

A multiplicative Kantian equilibrium of the game in normal form among a continuum of players with payoff functions $V_{(p,u)}$ is a strategy profile $\{a(p,u)\}$ and a coverage rate $\bar{a} = \int a(p,u)dQ(p,u)$ such that no player would prefer to re-scale the profile by any non-negative factor. We truncate the re-scaled probabilities (strategies) so that they do not exceed one upon re-scaling. The profile $\{a(p,u)\}$ is what the statistician observes: she does not observe the realization of the random variable L. We denote the profile of vaccination strategies of *individuals,* who know their realization of ε, by $\{\alpha^K(p,u,\varepsilon)\}$, whose mean is $\bar{\alpha}^K = \int \alpha^K(p,u,\varepsilon)dL(\varepsilon)dQ(p,u)$ We call the α^K- profile a *Kantian equilibrium of the vaccination game after the random variate L is realized* if no player (p,u,ε) would like to re-scale the entire profile by any non-negative factor ρ .

We must distinguish between the equilibrium after the random variable L has assigned a value ε to every player, and what the statistician observes, not knowing the realization of L. Since at the observed equilibrium there will be players with all values of ε at a given (p,u) in the support of Q, and these players will have different strategies $\alpha^K(p,u,\varepsilon)$, what the statistician will observe is that the $(p,u)-$ tranche is playing a mixed strategy:

$$a^K(p,u) = \int \alpha^K(p,u,\varepsilon)dL(\varepsilon)\,. \tag{2.7}$$

Note that $\bar{a}^K = \int a^K(p,u)dQ(p,u) = \bar{\alpha}^K$, because we have already integrated over both (p,u) and ε in the definition of $\bar{\alpha}^K$. $\bar{a}^K$ or $\bar{\alpha}^K$ is the coverage rate in the population, observed by the statistician.

Recall the definition of $\overline{\rho \circ a}$ from page 3. The expected utility of the parent in a profile re-scaled by the factor ρ is given by:

$$\tilde{V}_{(p,u,\varepsilon)}(\alpha,\bar{a},\varepsilon;\rho) =$$

$$\begin{cases} \tilde{V}^{+}_{(p,u)}(\alpha,\bar{a},\varepsilon;\rho) := \rho\alpha\overbrace{\big((1-p)\cdot 1 + pu + \varepsilon\big)}^{\text{ex utility if vacc}} + (1-\rho\alpha)\overbrace{\big(\pi(\overline{\rho \circ a})(1-p_0) + (1-\pi(\overline{\rho \circ a})\big)}^{\text{ex utility if not vacc}} & \text{if} \quad \rho > 1 \\ \tilde{V}^{-}_{(p,u)}(\alpha,\bar{a},\varepsilon;\rho) := \rho\alpha\overbrace{\big((1-p)\cdot 1 + pu + \varepsilon\big)}^{\text{ex utility if vacc}} + (1-\rho\alpha)\overbrace{\Big(\pi(\rho\bar{a})(1-p_0) + \big(1-\pi(\rho\bar{a})\big)\Big)}^{\text{ex utility if not vacc}} & \text{if} \quad \rho \le 1 \end{cases} \tag{2.8}$$

Note the function $\tilde{V}_{(p,u,\varepsilon)}$ is continuous, since $\lim\limits_{\rho\downarrow 1} V^{+}_{(p,u,\varepsilon)}(\alpha,\bar{a},\varepsilon;\rho) = V^{-}_{(p,u,\varepsilon)}(\alpha,\bar{a},\varepsilon;1)$, although it is not differentiable at $\rho = 1$. Furthermore, we have that for $\rho\epsilon[1,1/\alpha(p,u)]$:

$$\tilde{V}^{+}_{(p,u)}(\alpha,\bar{a},\varepsilon;\rho) - \tilde{V}^{-}_{(p,u)}(\alpha,\bar{a},\varepsilon;\rho) = (1-\rho\alpha)p_0\big(\pi(\rho\bar{a}) - \pi(\overline{\rho \circ a})\big) < 0, \tag{2.9}$$

because $\rho\bar{a} > \overline{\rho \circ a}$ on this interval.

Define the strategy profile after the random variate *L* has been realized:

$$\alpha^K(p,u,\varepsilon) = \begin{cases} \dfrac{-p_0\pi'(\bar{a}^K)\bar{a}^K}{(1-u)\big(p-g(u,\bar{a}^K)\big)-\varepsilon-p_0\pi'(\bar{a}^K)\bar{a}^K}, & \text{if } \varepsilon < (1-\mathrm{u})\big(p-g(u,\bar{a})\big) \\ 1 & \text{if } \varepsilon \ge (1-\mathrm{u})\big(p-g(u,\bar{a})\big) \end{cases}. \tag{2.10}$$

Note that on the first branch of this strategy profile (relatively small values of ε), the proposed strategy (probability) is less than one.

We have:

Proposition 3 *If* $\pi(\cdot)$ *is a decreasing, convex, twice-differentiable function on* [0,1], *then a multiplicative Kantian equilibrium exists and is given by the strategy profile defined in* (2.10).

Proof: The proof is provided in Appendix B.□□

From the proof of Proposition 3 we obtain the following condition for the existence of a Kantian equilibrium:

$$\bar{a}^K = \iint_{-\infty}^{(1-u)\left(p-g(u,\bar{a}^K)\right)} \frac{-p_0\pi'(\bar{a}^K)\bar{a}^K}{(1-u)\left(p-g(u,\bar{a}^K)\right)-\varepsilon-p_0\pi'(\bar{a}^K)\bar{a}^K} dL(\varepsilon)dQ(p,u) +$$

$$\int \left[1 - L\left((1-u)\left(p - g(u,\bar{a}^K)\right)\right)\right] dQ(p,u), \tag{2.11}$$

which is an equation in the single unknown $\bar{a}^K$. We solve for the Kantian equilibrium (2.10) by first solving (2.11) for $\bar{a}^K$ and then computing the equilibrium strategy profile from (2.10).

2.4 Comparison of Kantian and Nash vaccination equilibria

We noted in Section 1 that the vaccination game is a monotone increasing game. (Just check in equation (2.1) that $V_{(p,u)}$ is an increasing function of $\bar{a}$.) This is the mathematical consequence of the positive externality of individual vaccination. It follows that the Nash equilibrium of the game will suffer from the free-rider problem, but the multiplicative Kantian equilibrium will be Pareto efficient. Intuitively, people will vaccinate 'too little' in the Nash equilibrium. The precise consequence is this:

Proposition 4 $\bar{a}^K > \bar{a}^N$.

The equilibrium coverage rate is greater in Kantian equilibrium than in Nash equilibrium.

Proof:

Suppose to the contrary that $\bar{a}^N \geq \bar{a}^K$. Then $g(u,\bar{a}^K) \geq g(u,\bar{a}^N)$ and this implies that the second term in the r.h.s. of equation (2.11) is greater than the r.h.s. of equation (2.4). A fortiori, $\bar{a}^K > \bar{a}^N$ because the first term on the r.h.s. of equation (2.11) is positive. This contradiction proves the claim. n

In fact, we can say more. Note that although we have defined a parental type as an ordered pair of traits/beliefs (p,u), in fact the population profile of traits can be more parsimoniously written as depending only on the single variable $w = p(1-u)$. For we can write the Nash and Kantian equilibrium policies, from equations (2.5) and (2.11) respectively as:

$$\tilde{a}^N(w) = 1 - L\left(p(1-u) - p_0\pi(\bar{a}^N)\right) = 1 - L\left(w - p_0\pi(\bar{a}^N)\right). \tag{2.12}$$

and:

$$\tilde{a}^K = \int_{-\infty}^{w-p_0\pi(\bar{a}^K)} \frac{-p_0\pi'(\bar{a}^K)\bar{a}^K}{w-p_0\pi(\bar{a}^K)-\varepsilon-p_0\pi'(\bar{a}^K)\bar{a}^K} dL(\varepsilon) + 1 - L\big(w - p_0\pi(\bar{a}^K)\big).\text{[4]} \tag{2.13}$$

Since the domain of (p,u) is the unit square, the domain of *w* is [0,1]. We can plot the difference of the two equilibrium profiles

$$\Delta\tilde{a}(w) = \tilde{a}^K(w) - \tilde{a}^N(w). \tag{2.14}$$

See

Figure 2a in Section 5 below. When *w* is small then either *p* is small or *u* is close to one, or both, so the parent either believes that the probability of a severe side effect from vaccination is small, or if the side effect occurs, it is not severe (*u* close to one means the health status of a child with the side effect is close to full health). So parents with *w* close to zero will be likely to vaccinate according to our model and parents with *w* close to one will be likely not to vaccinate.

3 Producing the data

The data we require to compute the Kantian and Nash equilibria for a society are p_0, the distribution Q of (p,u), and the function π(·). We describe the choice of the logistic variate *L* below. We have administered the survey to adults aged 20 to 45 in the US, the UK, Germany, France, Canada, and Mexico. The survey is presented in Section III of the Online Appendix.

Table 2 shows some descriptive statistics.

Table 2. Descriptive statistics of the country surveys. Rows represent, respectively, the mean average among respondents, the percentage of females among respondents, the percentage of recent parents (those who had a child after in 2011 or later), and the percentage of those who vaccinated their child.

	Canada	France	Germany	UK	US	Mexico
Mean age	35.3	34.7	33.1	33.9	32.2	31.3
Female %	49.1	53.5	54.3	52.3	60.3	55.6
Parents since 2011 %	32.3	56.1*	33.8	41	32.8	56.5

[4] Note we have written the coverage rates in equations (2.12) and (2.13) as $\bar{a}^N$ and $\bar{a}^K$. We could have written these as $\bar{\tilde{a}}^N$ and $\bar{\tilde{a}}^K$. The coverage rates will be the same whether we integrate $dQ(p,u)$ or $d\tilde{Q}(w)$, where $\tilde{Q}(\cdot)$ is the distribution function of *w* induced by *Q*.

Measles vaccine %	88.9	90.3	89.3	89.2	82.9	96.8
N	1052	1188	1146	1054	1210	1063

* In the French survey the question asked was "Do you have a child born in or before 2018?".

The probabilities p and p_0 representing the individual's beliefs are ascertained in a standard way in the questionnaire.

We estimate u by presenting the respondent with a series of binary choices over pairs of lotteries. This technique allows us to place the respondent's value of u in a relatively small interval within $[0,1]$. The method assumes the individual is an expected utility maximizer.[5] We pose the question:

- In the following scenario, would you prefer event A or event B:
  A. For your child to have a bad side effect from a measles vaccination, or
  B. For your child to face an unrelated risk in which he/she has a 99% chance of being healthy, and a 1% chance of dying.

Suppose the respondent answers B. If utility is normalized as in Table 1, then we conclude that $(0.99 \times 1 + 0.01 \times 0) = 0.99 > u$. Next, we ask:

- In the following scenario, would you prefer event A or event B:
  A. For your child to have a bad side effect from a measles vaccination, or
  B. For your child to face an unrelated risk in which he/she has a 95% chance of being healthy and a 5% chance of dying.

Suppose the respondent answers A. Then we conclude that $u > (0.95 \times 1 + 0.05 \times 0) = 0.95$, and hence we know that $u \in (0.95, 0.99)$. We assign this respondent a value of u chosen randomly from this interval. Thus, we ascertain the respondent's value of u by posing a series of such questions about lottery choice.

We then fit a bivariate Beta distribution defined on $[0,1]^2$ to the respondents' values of (p, u). The Beta distribution is calculated knowing the observed means and variances of p and u, and their covariance. Table 3 presents these data for our six countries.

[5] See Holt and Laury (2002) for a description of this approach.

Table 3. Data from the country surveys. The probabilities p, the probability of side effects from vaccination, and p_0, the probability of death or severe disability conditional upon contracting measles, are ascertained in a standard way in the questionnaire. The utility $u \in (0,1)$ from the possible side effects is estimated from a series of binary choices over pairs of lotteries. The utility of a healthy child is normalized to 1.

	Mean p	Var p	Mean u	Var u	Cov(p,u)	Median p_0
US	0.048	0.026	0.851	0.061	-0.009	0.003
UK	0.020	0.009	0.891	0.043	-0.004	0.002
Germany	0.020	0.010	0.863	0.054	-0.004	0.002
France	0.022	0.011	0.743	0.094	-0.002	0.003
Canada	0.017	0.052	0.874	0.052	-0.0009	0.001
Mexico	0.035	0.015	0.774	0.068	-0.0005	0.005

We choose the common value of p_0 from the survey to be the median response to the appropriate question on the survey. The median is a better choice than the mean value, as the latter is distorted by several very high and unreasonable values for p_0.[6]

We comment on the value of p_0, the median value of respondents' opinions on the probability of dying from a measles infection. Dabbagh, Laws et alii (2018, Table 1) report that in 2017, for the continent of Europe the actual value is $p_0 = 0.004 = 0.4\%$, slightly larger than the median respondent's opinion. Unfortunately, this article does not present the value for the United States. But for Africa, the reported value of p_0 has a point estimate of 0.66, and the lower-bound estimate in the 95% confidence interval is 0.31. Measles can be a deadly disease if medical care is poor.

The most severe side effect of MMR (measles, mumps, rubella) vaccination is aseptic meningitis, which occurs in 1 in 10 million cases. The probabilities p that respondents to our survey give are greater than this by four orders of magnitude; however, from the values of u respondents provide, they are on average viewing side effects as not terribly severe (a value of $u = 0.85$ says that good health is 15% reduced by the side effect). The possibly bad outcomes of measles are considerably worse, and include, besides death, anaphylaxis, febrile seizures, thrombocytopenic purpura and encephalitis (see Strebel and Orenstein,

[6] However, we report the mean values of p and u because these are used to fit the Beta distribution Q to the data.

2019, which also gives the probabilities). The anti-vaccination movement is often motivated by fears that vaccination may cause autism, which were falsely aroused in a 1998 article published in *Lancet,* later retracted by *Lancet* in 2010.

We use the following parametric form for the probability function:

$$\pi(\bar{a}) = (1 - \bar{a})^{\gamma}, \quad (3.1)$$

where $\bar{a}$ is the observed measles vaccination coverage rate for the country. We chose the parameterization (3.1) as possibly the simplest functional form that gives a decreasing function passing through the points (0,1) and (1,0). In Appendix B, we describe the precise definition of the function π and how we estimate γ. For Canada and the United States, we estimate γ =3.1. For the UK, France, Germany and Sweden, we estimate $\gamma = 1.995$.[7] We split our set of countries in two because the number of cases of measles in the last five years in Europe has been an order of magnitude larger than in North America (Canada and the US), despite the higher coverage rates enjoyed by the European countries. We presume the infection process therefore differs between recent European experience and the North American, justifying different values of γ in equation (3.1).[8]

At the WHO-reported[9] coverage rate of 0.916 for the US, the probability that an unvaccinated child in a given cohort in the US contracts measles before the age of five, defined as the number of measles cases in her birth cohort divided by the number of unvaccinated children in her cohort, is 4.5×10^{-4}, or about 0.045%.

[7] We had planned to include Sweden in our sample of countries, and so included it in the estimation of γ. Unfortunately, doing so was eventually not possible. Estimating the European value of γ without Sweden gives a value of 2.007. Based on the small difference between this value and 1.995, we elected not to re-run all the equilibrium calculations for the UK, France and Germany with $\gamma = 2.007$, a costly procedure.

[8] Without morbidity data for Mexico, we use the European value of $\gamma = 1.995$. Nevertheless, results do not change significantly when using the North American value.

[9] Data source https://apps.who.int/immunization_monitoring/globalsummary/.

The last year measles was endemic in the United States was 2000.[10] The aforementioned WHO data set reports that in 2019, measles was endemic in Germany and France. It is probably also endemic in Mexico, although the data are incomplete. We cannot use the SIR model to compute the probability of contracting measles because this model is not applicable to analyzing very small occurrences of the disease that are quickly stamped out[11]. In any case, the SIR model will not give us a probability as a function of the coverage rate only: in that model, the probability that a susceptible individual contracts the disease is a function of two numbers –for instance, the fraction of susceptible (uninoculated) individuals (S), and the fraction of recovered individuals (R).

Our definition of the function π as the probability that a child who is unvaccinated contracts measles by the age of five is meant to model the relevant probability that a parent needs in order to decide whether or not to vaccinate her child.

Table 4. Coverage rates for measles, five-year average, according to the World Health Organization.

Country	US	UK	Germany	France	Canada	Mexico
Coverage	91.6%	92.%	97.%	90.2%	89.6%	86%

It is important to note that measles vaccination in our six countries is, or was until recently, *de jure* voluntary. In the United States, there is no federal law requiring children be vaccinated –such laws are left to the states. All 50 states require children be vaccinated against measles before attending childcare or public school; however, all states permit exemptions for medical, religious, or reasons of conscience, and the standards are not strict. In Canada, vaccination policies are taken at the provincial level. Only three provinces (Ontario, New Brunswick and Manitoba) have legislated requirements; however, exemptions are granted on medical or religious grounds, or simply out of conscience in these provinces. In the UK,

[10] A contagious disease is endemic if an outbreak induces a sequence of contagion that does not terminate within a year.

[11] A useful description of the SIR model is found in Avery, Bossert et al (2020).

childhood vaccination is not mandatory. In Germany, a federal law now requires measles vaccination, but only since March 1, 2020. In our German survey, we asked parents whether they vaccinated or did not vaccinate their child prior to that date. In France, vaccination was only recommended prior to 2018, and in the French questionnaire, we asked parents for the vaccination status of their child prior to 2018. Mexico has no law requiring vaccination.

4 Estimation procedure

We wish to decide whether the Nash model or the Kantian model provides a better explanation of observed vaccination behavior in a country. We have samples of roughly 1000 (*N*) respondents for each country. Each respondent is characterized by a triple (p, u, v) where $(p, u) \in [0,1]^2$ is the vector of respondent traits and $v \in \{0,1\}$ indicates that the respondent did (1) or did not (0) vaccinate her child. We call $\boldsymbol{v}^{obs,s^0} = (v^1, \dots, v^N)$ the *observation* or *observed vaccination behavior* of the original sample. The superscript s^0 refers to the original survey sample for the country.

There are three sources of randomness in our models. First, there is a logistic variate L, 99% of whose mass lies on the positive real line (more below). Each parent who chooses to vaccinate draws a realization of this variate i.i.d. across individuals, which is interpreted as a (usually) positive saltus in utility that the parent enjoys if she vaccinates her child (see (2.1)). Secondly, since the equilibrium strategies observed by the statistician in both the Nash and Kantian model are mixed strategies, there is a random process which must determine whether a player with an equilibrium strategy $a \in (0,1)$ chooses $v = 0$ or 1. Third, there is a 'trembling hand' introduced below: with some probability q each player, when choosing the action *v*, misreads the coin flip that determines what her behavior should be. (These trembles will be i.i.d.) The purpose of the first and third sources of randomness is to make the models more realistic, so as to achieve a better fit to the observed vaccination behavior, and to guarantee that the Nash and Kant equilibrium strategies are all strictly mixed strategies (lie in the open interval (0,1)). The second source is due to the mixed-strategy character of the equilibria.

4.1 The logistic variate L

It is useful for computation to have the support of L be the entire real line: this guarantees that all equilibrium strategies, Nash and Kant, are in the open interval (0,1). This motivates our choice of a logistic distribution. See equations (2.5) and (2.13), which guarantee that the probabilities of vaccination are never zero or one when L's support is $\mathbb{R}$. We shall determine L by a single parameter, its mean value μ. The logistic variate is in fact characterized by two parameters, denoted (μ, β) . Denote by $L^{(\mu,\beta)}$ the c.d.f. of the logistic with parameters(μ, β). Given μ, we choose β so that:

$$L^{(\mu,\beta)}(0) = 0.01; \tag{4.1}$$

that is, 99% of L's mass is on the positive real line. Hence L is chosen from a single parameter family, where the parameter is μ. We chose $\mu = 0.003$ and performed a robustness check by running the program for other values of μ. (See Section II of the Online Appendix.)[12]

4.2 Nash and Kantian equilibria

We will perform the estimation procedure outlined in this section for a large number B of bootstrapped samples, obtained from the original survey sample s^0 by sampling from it with replacement. Let the size of the mother sample s^0 and of all the bootstraps for a particular country be *N.* Here we describe the estimation procedure using the mother sample s^0; the identical procedure will be carried out for every bootstrap sample s.

Given the sample s^0, we fit a bivariate Beta distribution Q^0 to the observed distribution of (p, u). μ is chosen to be a small positive number. For any choice of μ , the logistic distribution $L^{(\mu,\beta)}$ is determined, see (4.1). Given L and Q^0 we can compute the Nash and Kantian equilibria of the vaccination game observed by the statistician as described in Section 2. The Nash equilibrium is a profile of strategies

[12] We also tried to estimate μ as the value that maximized the average likelihood among all the bootstrapped samples. We generated 1000 bootstrapped samples and computed the likelihoods for each $\mu \in \{0.001, 0.002, \ldots, 0.008\}$. The average likelihood maximizers were not the same for the Nash and the Kantian equilibria, making the comparison ineffective. We opted then for choosing the value of μ that most frequently maximized the likelihood across samples and run a robustness check.

(probabilities of vaccinating) $a^N(p,u{:}\,s^0,\mu)$ and the Kantian equilibrium is a profile of strategies $a^K(p,u{:}\,s^0,\mu)$.

Given these two equilibria, we can compute *the log likelihood of the observed vaccination behavior* $\mathbf{v}^{obs,s^0}$. This is defined, for the Nash equilibrium, as:

$$\Phi^{Na}\left(s^0,\mu,\mathbf{v}^{obs,s^0}\right)=\sum_{\{(p,u)|v=1\}}\log a^N(p,u;s^0,\mu)+\sum_{\{(p,u)|v=0\}}\log\left(1-a^N(p,u;s^0,\mu)\right), \quad (4.2)$$

and for the Kantian equilibrium as:

$$\Phi^{Ka}\left(s^0,\mu,\mathbf{v}^{obs,s^0}\right)=\sum_{\{(p,u)|v=1\}}\log a^K(p,u;s^0,\mu)+\sum_{\{(p,u)|v=0\}}\log\left(1-a^K(p,u;s^0,\mu)\right), \quad (4.3)$$

where the original sample is the collection of triples $\{(p,u,v)\}$.

Since the strategies are all in the open interval $(0,1)$, the two log likelihood functions are well-defined. Because of precision problems in computation, we in fact encounter some zero values in the computation of $a^N(p,u)$. Rather than eliminating these respondents from the sample, we replace the zero values of $a^N(p,u)$ with $ra^K(p,u)$ where $r=\underset{\{(p,u)|a^N(p,u)>0\}}{\text{mean}}[a^N(p,u)/a^K(p,u)]$. It will turn out that $r<1$, because $a^N(p,u)<a^K(p,u)$ for all (p,u).

4.3 Analyzing the sample

Next, we ask: Could it be that $\mathbf{v}^{obs,s^0}$ can be explained as an outcome of Nash behavior, but amended by a trembling hand that causes each respondent to choose the opposite behavior from what the Nash coin-flip produces? Let's say the tremble occurs i.i.d. for each respondent with probability *q*. In this case, an agent (p,u) chooses to vaccinate $(v=1)$ with probability :

$$a^{*N}(p,u)=(1-q)a^N(p,u)+q\left(1-a^N(p,u)\right). \quad (4.4)$$

Suppose we run a large number, Λ, of trials with this model, all with the sample s^0. The only thing that differs across trials is the realization of the coin flips that implement the tremble: the expected value of the coin flip for an agent (p,u) is always given by $a^{*N}(p,u)$ in (4.4). Denote the index of the trial by *l*. Define:

$$\mathbf{1}_q^l=\{(p,u)|a^{*N}(p,u)\text{ coinflip }l\to 1\},\mathbf{0}_q^l=\{(p,u)|a^{*N}(p,u)\text{ coinflip }l\to 0\}$$

We ask: What log likelihood would *this* observed vaccination outcome have if we mistakenly thought the true Nash model (absent the coin-flip) were the correct model? That likelihood is given by:

$$\Psi(q,l;s^0,\mu)=\sum_{(p,u)\in\mathbf{1}_q^l}\log a^N(p,u;s^0,\mu)+\sum_{(p,u)\in\mathbf{0}_q^l}\log\left(1-a^N(p,u;s^0,\mu)\right). \quad (4.5)$$

We are taking the log likelihood of the observed behavior from the trembling-hand coin-flip experiment and evaluating it with respect to the *pure* Nash model, without the trembling hand.

Next, we want to compute the expected value of $\Psi(q,l)$ over $l=1,2,\dots,\Lambda$. We can write the expected log likelihood of the experiment as the number of trials Λ becomes large as:

$$M(q)\equiv\lim_{\Lambda\to\infty}\frac{1}{\Lambda}\sum_{l=1}^{\Lambda}\Psi(q,l)=\sum_{(p,u)}[a^{*N}(p,u)\log a^N(p,u)+\left(1-a^{*N}(p,u)\right)\ \log\left(1-a^N(p,u)\right)].$$

(4.6)

This is the key step. It's true because if we look at the sums in (4.5) over all *l*, by definition, a given (*p*,*u*) will lie in the set $\mathbf{1}_q^l$ for a fraction $a^{*N}(p,u)$ of the Λ trials, as Λ becomes large. And (p,u) will lie in $\mathbf{0}_q^l$ a fraction $1-a^{*N}(p,u)$ of the time.

Our strategy is to ask how large a tremble is needed to produce the log likelihood $\Phi^{Na}\left(s,\mu,\mathbf{v}^{obs,s^0}\right)$. Our claim is: the smaller the tremble needed to 'rationalize' the observed vaccination behavior, the better explanation the model provides of observed behavior. In other words, since we view the trembling hand as a device for inserting randomness into the Nash (or Kant) model, then the less randomness required to explain the observed behavior, the better the model's explanatory power.

Consequently, we wish to solve the following program for the tremble *q*:

$$\min_q\left(M(q)-\Phi^{Na}\left(s,\mu,\mathbf{v}^{obs,s^0}\right)\right)^2. \quad (4.7)$$

Note, from the definition (4.6), that $M(\cdot)$ is a linear function of q. So we can solve program (4.7)) by setting the derivative of the objective equal to zero. Compute from (4.6) that:

$$M'(q)=\sum_{\{(p,u)|0<a^N(p,u)<1\}}\left(1-2a^N(p,u)\right)\log\frac{a^N(p,u)}{1-a^N(p,u)} \quad (4.8)$$

Now the f.o.c. for program (4.7) is:

$$2\left(M(q)-\Phi^{Na}\left(s^0,\mu,v^{obs,s^0}\right)\right)M'(q)=0. \quad (4.9)$$

From (4.8), we see that generically, $M'(q) < 0$. (Each term in the sum in (4.8) is negative, except if $a^N = 1/2$.) Therefore, the solution of (4.9) requires:

$$M(q) = \Phi^{Na}\left(s^0, \mu, v^{obs,s^0}\right). \tag{4.10}$$

Recalling that M is linear, we easily solve (4.10) for q, giving:

$$q_\mu^{*\text{Nash}} = \frac{\Phi^{Na}\left(s,\mu,\mathbf{v}^{obs,s^0}\right) - \left[\sum_{\{(p,u)|0<a^N<1\}} a^N \log a^N + \sum_{\{(p,u)|0<a^N<1\}} (1-a^N)\ \log(1-a^N)\right]}{\sum_{\{(p,u)|0<a^N<1\}} (1-2a^N) \log\frac{a^N}{1-a^N}}. \tag{4.11}$$

Actually, this is the solution if the quantity on the r.h.s. of (4.11) lies in $[0,1]$. If the r.h.s. of (4.11) is greater than 1, then $q_\mu^{*N} = 1$ and if it is less than 0, then $q_\mu^{*N} = 0$. In other words, if the true solution of (4.9) were at a corner of [0,1], the f.o.c. (4.10) becomes an inequality.[13]

This completes the estimation procedure for the sample s^0. We repeat the estimation procedure for each of B=1200 bootstrap samples. Denote, for bootstrap sample s, the q*'s* defined in equation (4.11) as $q^{*J}(s)$, for J =Nash, Kant.

We finally define two functions for all bootstrap samples s:

$$\Delta(s) = q^{*K}(s) - q^{*N}(s) \text{ and } \Gamma(s) = \Phi^K\left(s, \mu, \mathbf{v}^{obs,s}\right) - \Phi^N\left(s, \mu, \mathbf{v}^{obs,s}\right), \tag{4.12}$$

and deduce statistics on Δ and Γ using the 1200 bootstrap samples. For instance, if we find that the mean of the distribution $\Delta(s)$ is negative and more than two standard deviations below zero, we will say that the Kant model provides a better explanation of vaccination behavior than the Nash model, at the 95% significance level. A similar inference would be drawn if $\Gamma(\mathrm{s})$ is positive and at least two standard deviations away from zero.

In Section II of the Online Appendix, we perform a robustness check by running the program for several values of μ.

---

[13] It turns out that for all our countries and samples, the numbers $q^{*J}(s)$, for J =Nash, Kant are in (0,1). This means that, at the optimal values of q, $M\left(q^{*J}(s)\right) = \Phi^J\left(s, \mu, \mathbf{v}^{obs,s}\right)$. We are able to adjust the tremble so that the expected value of the log likelihood of the trembling-hand model is precisely the *observed* log likelihood for that sample and model.

5 Major findings

We summarize the main findings of our analysis, namely the Nash and Kantian equilibrium strategy profiles for types (p, u) and their empirical fit with the observed vaccination behavior. The Online Appendix offers details on the survey, the bootstrap strategy, and the country-specific results, as well as brief historical discussions of measles vaccination in each country.

5.1 Equilibrium strategy profiles

For all countries, we find that the profile of Kantian equilibrium strategies dominates the profile of Nash equilibrium strategies: that is, for all types (p, u), $a^K(p, u) > a^N(p, u)$ or, equivalently, for all w, $\tilde{a}^K(w) > \tilde{a}^N(w)$. This is illustrated in two different spaces in

Figure *2*. Kantians always vaccinate with higher probability than Nashers. The continuous functions $\Delta\tilde{a}(w)$ are graphed in

Figure *2*a (recall the definitions in equations (2.12) – (2.14)). The observed values of w in any country sample comprise a set of approximately 1000 values, which will lie along these curves.

Figure *2*b presents the graphs of the actual equilibrium profiles in the space $\left(a^N(p, u), a^K(p, u)\right)$.

From

Figure *2*a, note that the differences between the Nash and Kantian strategies are greatest for Mexico: this is verified for the empirical distributions in the Mexican graph in

Figure *2*b. Contrast Mexico with Canada. We see from

Figure *2*a that the differences $\Delta\tilde{a}(w)$ are very small in Canada: this is verified in

Figure *2*b, where we see that observed strategy pairs are very close to (but lie above) the 45º line. We emphasize that the graphs in

Figure *2*a are derived from the estimated beta-distributions of types Q in the six countries.

From

Figure *2*b, it appears that the Kant and Nash equilibrium probabilities occur densely in the unit interval. That is, if our data (p,u) were dense in the unit square, equilibrium probabilities would likewise be dense in [0,1].

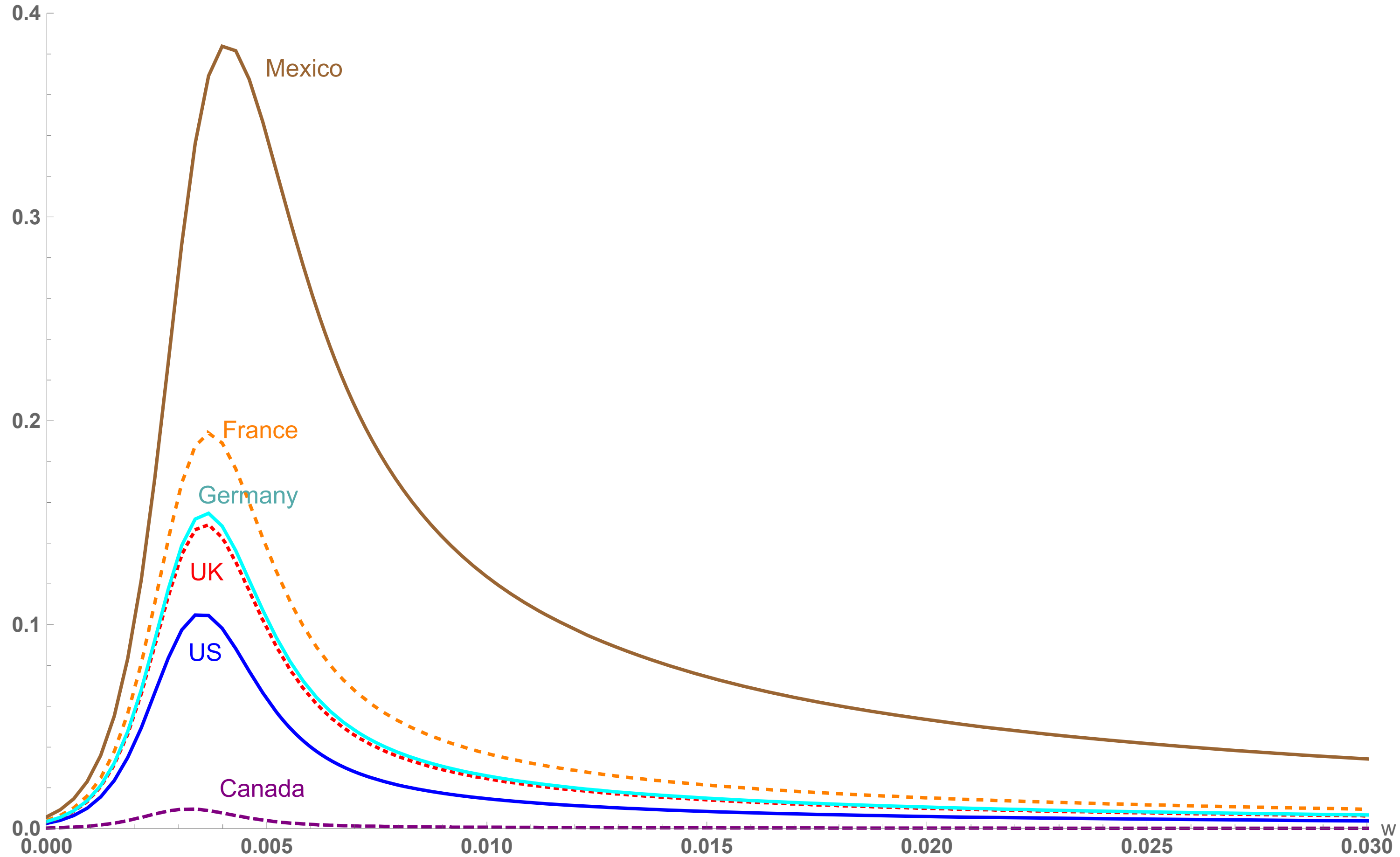

Figure 2a. The difference between the Kantian and the Nash equilibrium strategy profiles $\Delta\tilde{a}(w)$ across types of agents. The horizontal axis, $w = p(1-u)$, is the single variable that characterizes an agent's beliefs.

.

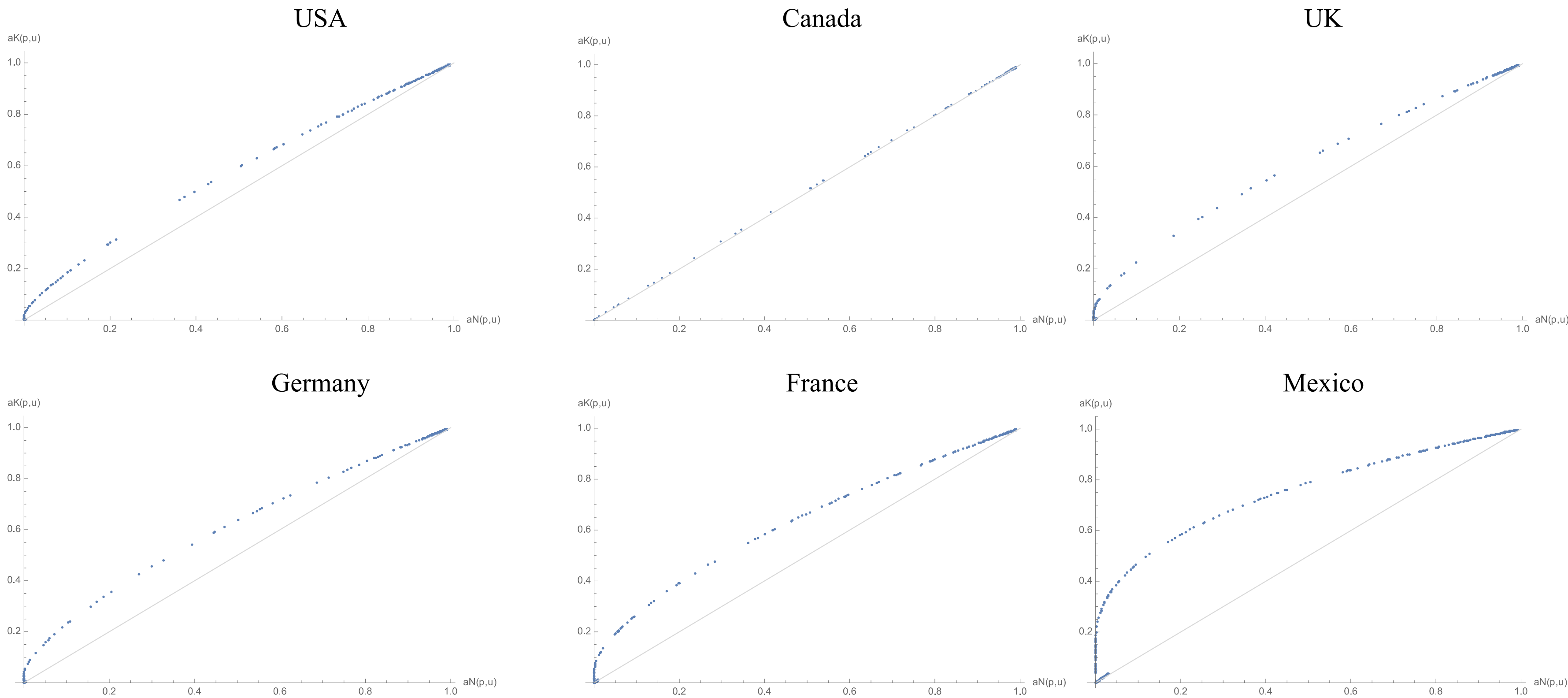

Figure 2b. Equilibrium strategy profiles from the original survey data. Kantian vs. Nash strategies for the observed strategy pairs $(\tilde{a}^N(w), \tilde{a}^K(w))$. Kant and Nash equilibrium probabilities appear to occur densely in the unit interval.

Figure 2. Kantian vs. Nash strategy profiles. We observe that Kantians always vaccinate with higher probability than Nashers, and that the differences between the Nash and Kantian strategies vary across countries, being greatest for Mexico and small for Canada. The observed values of w in any country sample comprise a set of approximately 1000 values.

5.2 Empirical fit

Our estimation procedure shows that the optimal tremble for the Kantian model, over all bootstrap samples, is significantly less than the optimal tremble for the Nash model. See Figure 3. In all six countries, the difference of the optimal trembles ($q^K - q^N$) is significantly less than zero at the 99.9% significance level (that is, $\Delta(s) < 0$). Our interpretation of this fact is that the Kantian model provides a significantly better explanation of vaccination behavior than the Nash model, as we discussed in Section 4. In addition, in all six countries, the difference of the log likelihood functions ($\Gamma(s) = \Phi^K(s) - \Phi^N(s)$) is significantly greater than zero at the 99.9% significance level. See Figure 4. The interpretation is, again, that the Kantian model provides a significantly better explanation of vaccination behavior than the Nash model. The significance levels are reported in the Online Appendix.

In figures Figure *3* and Figure *4*, we plot over each histogram the graph of the density function of the normal distribution with mean and standard deviation of the histogram, over the interval $\pm$ three standard deviations from the mean, verifying our claim concerning significance levels.

Rather than looking at the differences, Figure 5 gives the histogram over all bootstrap samples of the values of the optimal tremble for the Nash and Kant model. These differ somewhat across countries. Similarly, Figure 6 presents the histograms for the Nash and Kant log likelihood functions for all countries.

Finally, Figure 7 gives the histogram of the absolute value of the difference between the 'observed' coverage rate and the Nash (or Kant) equilibrium coverage rates at the optimal trembles, averaged over all samples. These differences are quite large. But recall that our estimation strategy is to minimize the tremble, not the difference between the equilibrium coverage rate and the observed coverage rate.

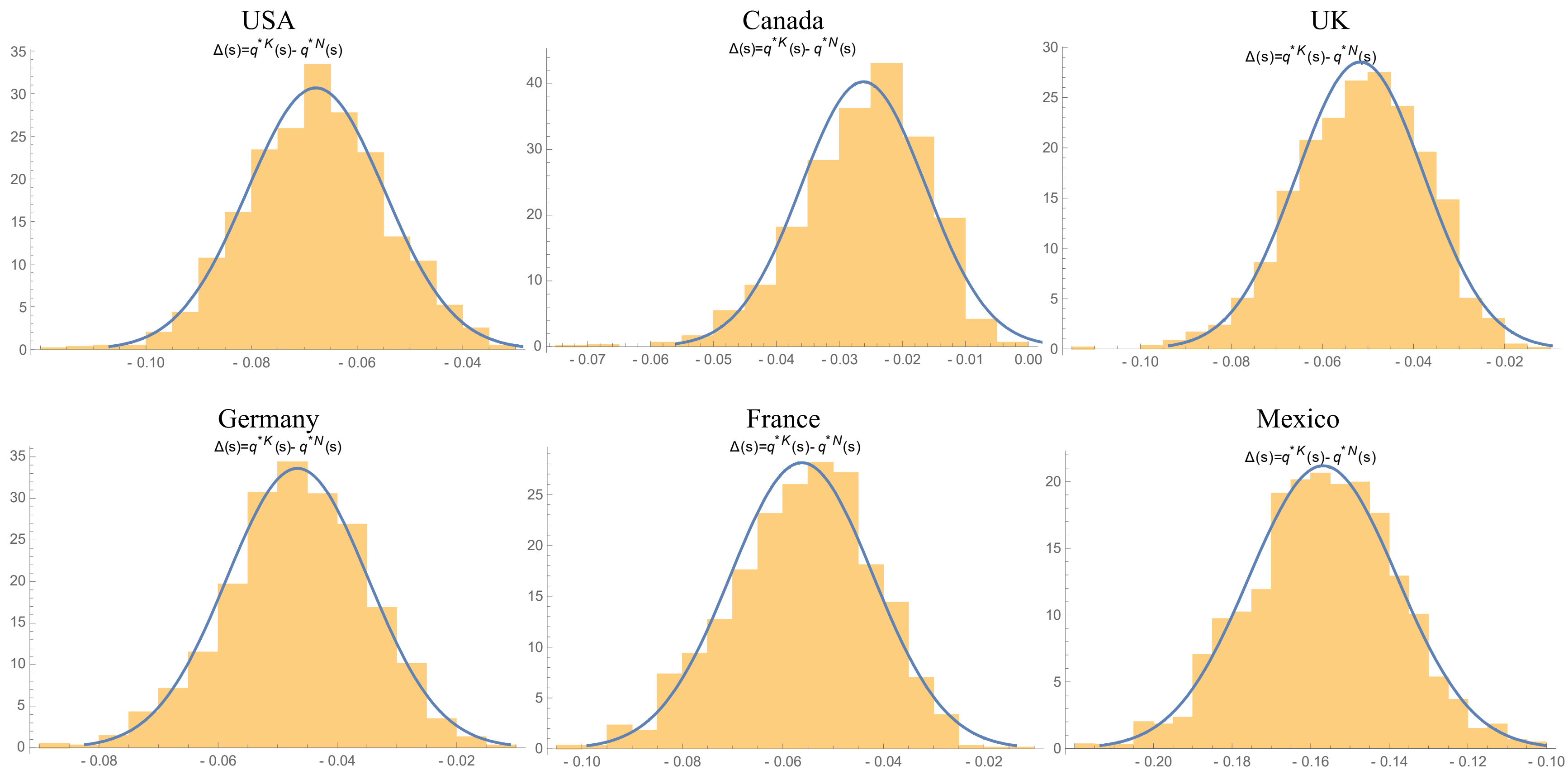

Figure 3. Probability density histogram of the differences between the optimal trembles ($\Delta = q^{*K} - q^{*N}$), and the PDF of a Normal distribution $N(m, \sigma)$ with $m =$ Mean (Δ) and $\sigma =$ StaDev (Δ), truncated at three standard deviations from the mean. The graphs show that the difference of the optimal trembles is significantly less than zero at the 99.9% significance level, supporting the inference that the Kantian model provides a significantly better explanation of vaccination behavior than the Nash model.

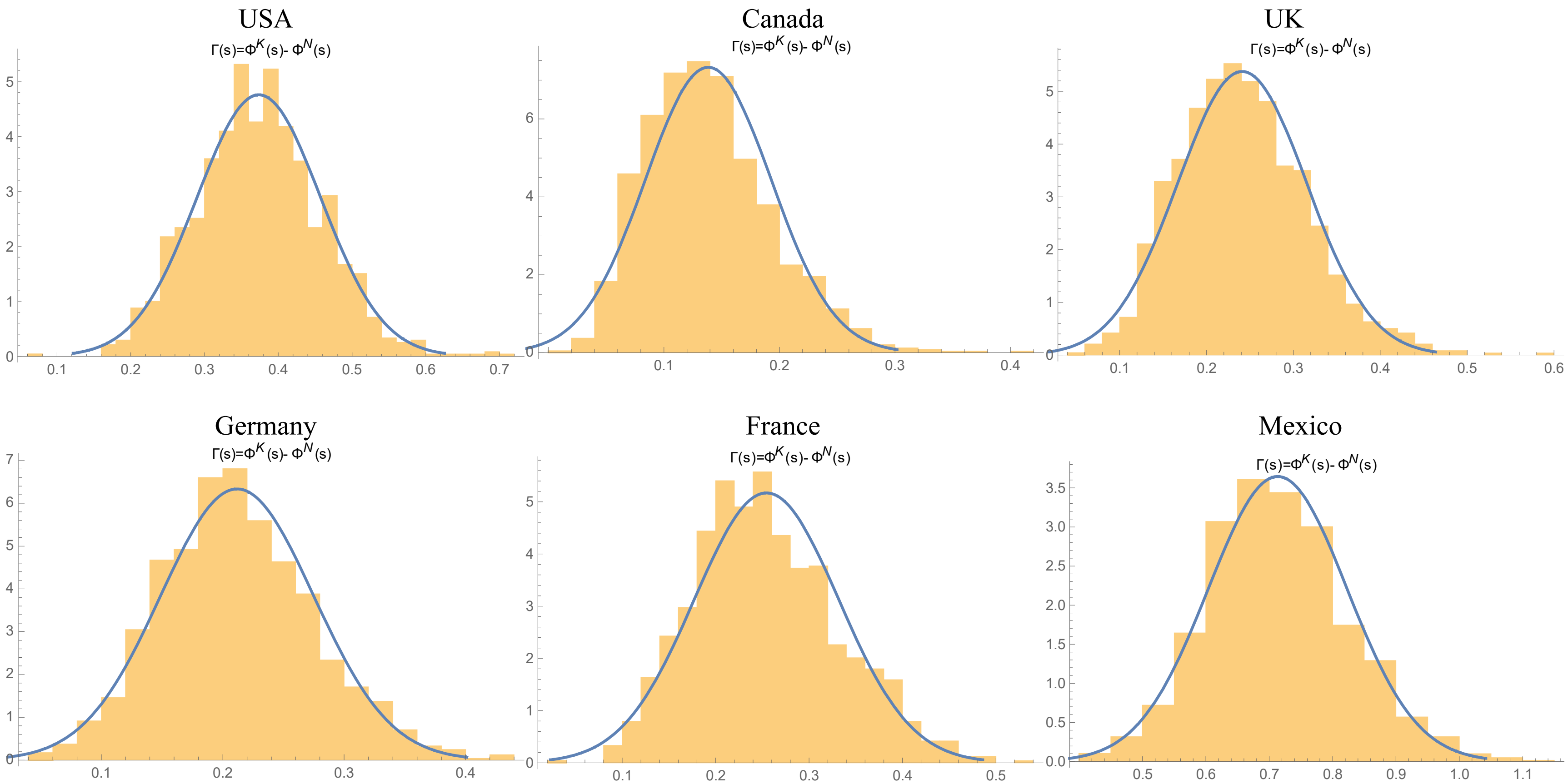

Figure 4. Probability density histogram of the difference of the per capita log-likelihoods ($\Gamma = \Phi^K - \Phi^N$), and the PDF of a normal distribution $N(m,\sigma)$ with $m =$ Mean (Γ) and $\sigma =$StaDev (Γ), truncated at three standard deviations from the mean. The graphs show that the difference of the log-likelihood functions is significantly greater than zero at the 99.9% significance level, supporting the inference that the Kantian model provides a significantly better explanation of vaccination behavior than the Nash model.

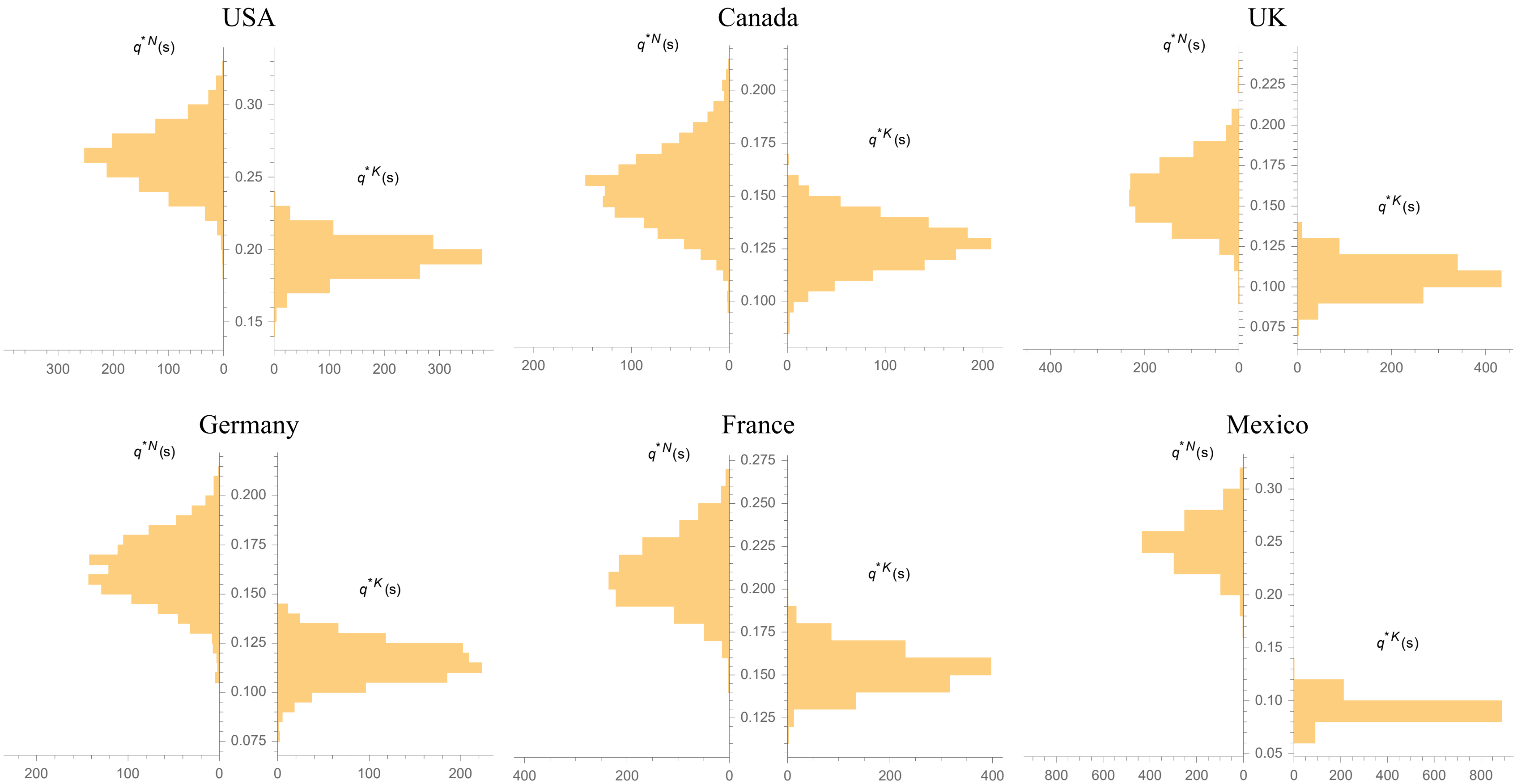

Figure 5. Histograms over all bootstrap samples of the values of the optimal tremble for the Nash (q^{*N}) and Kantian (q^{*K}) equilibrium.

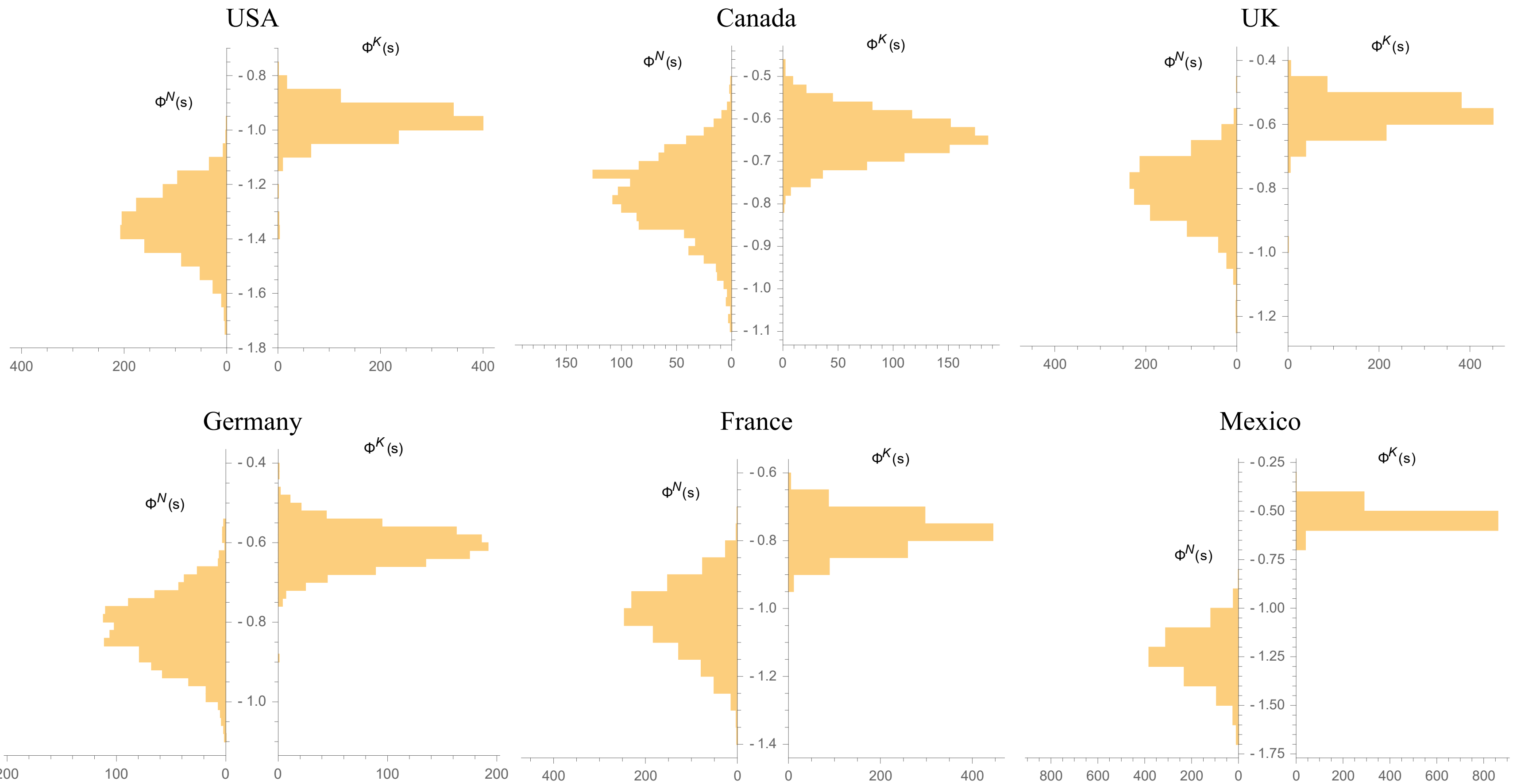

Figure 6. Histograms of the Nash (Φ^N) and Kantian (Φ^K) log-likelihood functions.

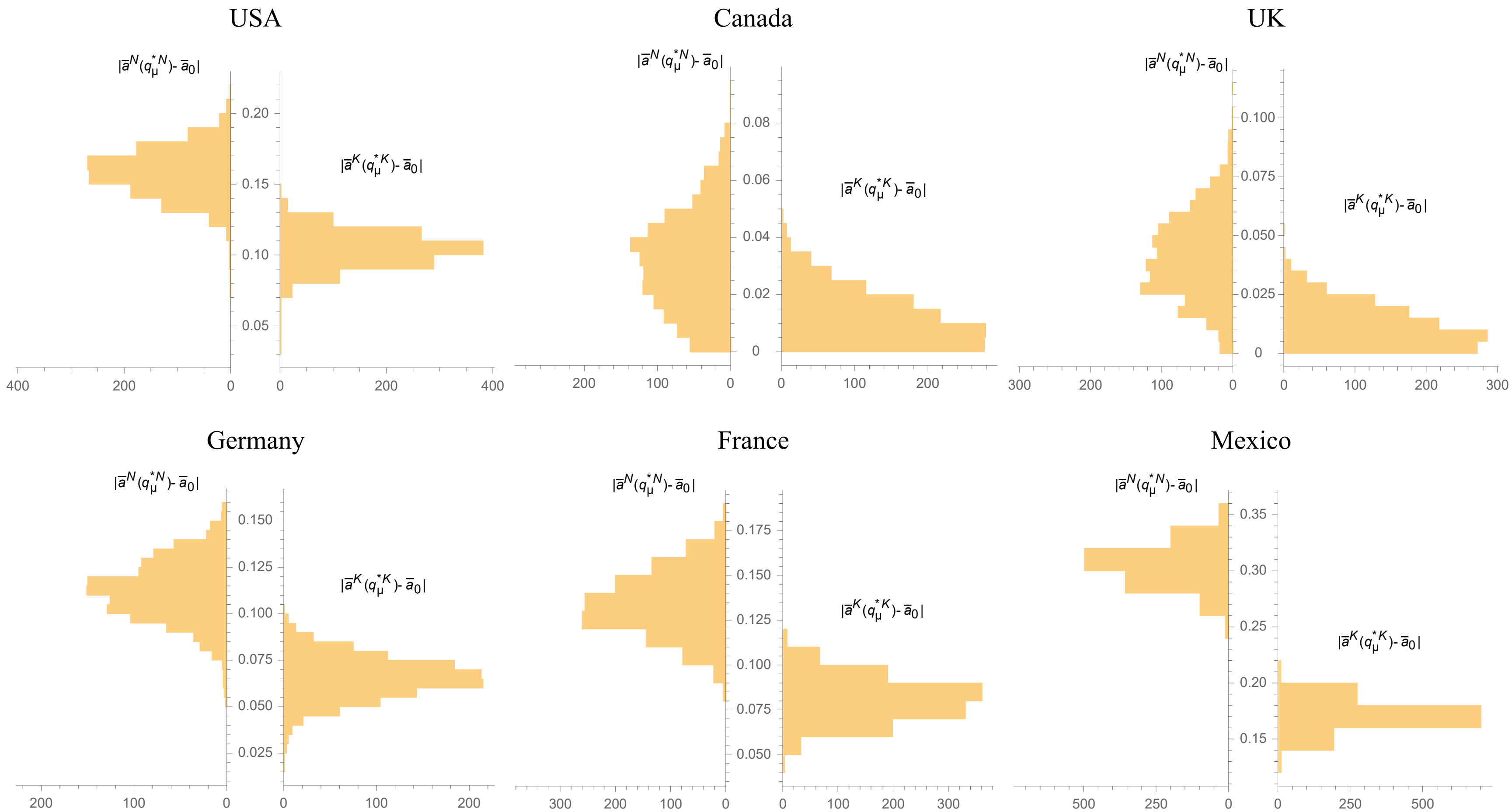

Figure 7. Histograms of the difference between the 'observed' coverage rate at the survey sample ($\bar{a}_0$) and the Nash and Kantian equilibrium coverage rates ($\bar{a}^J, J = N, K$) at the 'optimal' tremble, averaged over all samples.

6 Why does the Kantian model of vaccination give a superior explanation?

We propose that the reason the Kant model is superior to the Nash model is that it gives uniformly higher probabilities of vaccination than the latter, and the coverage rates in the Kantian equilibria are closer to the observed coverage rates in our samples than the Nash coverage rates. See the six panels in Figure 8, where we plot the equilibrium coverage rates in the Nash and Kant equilibria of our 1200 bootstrap samples, in comparison to the observed coverage rate of sample s^0 (for each country). We believe that actual vaccination behavior is more pervasive than the Nash model predicts because in all countries in our sample there is a social norm to vaccinate. The social norm shares with the Kantian equilibrium the property of inducing more pervasive vaccination behavior than Nash equilibrium predicts. Even though the Kantian equilibrium strategies appear to be not much larger than the Nash probabilities for many types, the fact that they are *always* larger than the latter, on the space of types, makes the likelihood of the Kantian equilibrium significantly greater than the likelihood of the Nash equilibrium. (See our major finding in Section 5.1 .). It is commonly believed that social norms are alternatives to Nash equilibrium, but the action of the norm is not well theorized. Kantian equilibrium gives a clean micro-foundation of such a social norm.[21] One virtue of our approach is that we can run a horse race between Nash and Kant because the two concepts use the same data and both are precisely defined.

Why has a social norm to vaccinate developed in all these countries? Because, we conjecture, of the positive externality from individual vaccination. We pointed out in Section 1 that the Nash equilibrium of the vaccination game will always be Pareto inefficient (Proposition 1), because of the positive externality associated with vaccination. Kantian optimization is *one way* to repair this inefficiency: *another* way to do so is through the development of a social norm to vaccinate. We are suggesting that the latter has indeed occurred in many societies, and this is why the Kantian equilibrium gives a better explanation of vaccination

[21] Thus, if we had a precise, general formulation of what a social norm is we could skip the Kantian step.

behavior than the Nash equilibrium. One might say that Kantian equilibrium is one precise prediction of what a social norm will produce.

We might point out that a similar explanation applies to explaining the payment of income taxes, which, in all highly developed countries is more pervasive than Nash equilibrium predicts. There is a literature showing that in the United States, given the penalties associated with tax evasion and the probability of being caught for doing it, evasion should be significantly higher than it is, assuming reasonable attitudes towards risk (see Alm (2019) and the many references therein). The reason is that there is a social norm to pay one's taxes. And this norm has developed precisely because the 'tax game' is a monotone increasing game: the taxes you and others pay make me better off, because they finance public goods and social insurance from which I benefit.[22]

To test this conjecture, we administered a second survey on the motivation to vaccinate, which was inadequately covered in our first survey, in the US and France, two countries that have significant anti-vax movements. We report the key findings of both surveys in the next three tables. In our follow-up survey we received 1243 responses from Americans and 1490 responses from French residents.[23] Herd immunity and vaccination behavior of others is clearly indicated to be encouraging rather than discouraging own vaccination, which is in line with Kantian optimization and a strong social norm, but not with Nash equilibrium. (A Nash optimizer will be discouraged to vaccinate her child if herd immunity is approached.) In a scenario of well-established herd immunity for a child illness, 68.4% of US respondents (57.5% of French ones) are either 'strongly encouraged or encouraged' to vaccinate their own child (Table 5), while only 6.24% (7.04%, resp.) are discouraged.

---

[22] Or perhaps the tax norm has developed for another reason: that people have an aversion to cheating, which may also be a social norm.

[23] These counts include all responses (including those who simply declined consent and ended the survey) but exclude any responses classified as "spam" by Qualtrics.

Table 5. Distribution of responses to the question Q4.6 "Imagine herd immunity is already well-established for a specific child illness because of a high vaccination rate. Would that encourage or discourage you from vaccinating your own child?" in the US and France surveys.

	US		France	
	Frequency	Percent	Frequency	Percent
Strongly encourage	475	43.58	372	28.48
Encourage	271	24.86	379	29.02
Leave unchanged	276	25.32	463	35.45
Discourage	41	3.76	58	4.44
Strongly discourage	27	2.48	34	2.60
N	1,090	100	1,306	100

The same reaction is observed to an individual act of vaccination. Learning that others have vaccinated their child 'strongly encourages or encourages' 61.4% of US respondents (and 45.9% of French ones, see Table 6). Likewise, own vaccination is expected to strongly encourage or encourage others' vaccination by 64.6% of US respondents (and 50.4% of French ones, see Table 7), suggesting that a parent believes a social norm is operative.

Table 6. Distribution of responses to the question Q4.3 "If you learn that others have vaccinated their child, would that encourage or discourage you to vaccinate your child?" in the US and France surveys.

	USA		France	
	Frequency	Percent	Frequency	Percent
Strongly encourage	390	35.78	236	18.02
Encourage	279	25.60	365	27.86
Leave unchanged	368	33.76	673	51.37
Discourage	25	2.29	15	1.15
Strongly discourage	28	2.57	21	1.60
N	1,090	100	1,310	100

Table 7. Distribution of responses to the question Q4.4 "When you vaccinate your child, would you expect others to be encouraged or discouraged by your action to also vaccinate their child?" in the US and France surveys.

	USA		France	
	Frequency	Percent	Frequency	Percent
Strongly encouraged	360	33.03	255	19.47
Encouraged	344	31.56	405	30.92
Leave unchanged	351	32.20	624	47.63
Discouraged	19	1.74	16	1.22
Strongly discouraged	16	1.47	10	0.76

N	1090	100	1310	100

For those respondents who have or would vaccinate their child, 73.5% of US respondents (and 72.6% of French ones) indicate that "Vaccination protects my child from disease" is a very important reason for that decision. Other reasons that relate to herd immunity or social norms are also deemed very important by large fractions of respondents, "Vaccination of my child contributes to herd immunity" by 54.9% of US respondents (46.2% of French ones) and "I vaccinate because other parents I know choose to vaccinate" by 35.4% of US respondents (20.4% of French ones). Conversely, for those respondents who have not or would not vaccinate their child, side effects and choice autonomy are deemed as very important more often than matters of herd immunity or social norms. This can be seen in the contrast between "There are possibly severe side effects to vaccination" (56.4% in the US, 61.6% in France) and "Vaccination should be a matter of free choice" (54.3% in the US, 58.0% in France) on the one side and on the other side "If vaccination coverage is already high in the community, my child will be safe without vaccination" (30.9% in the US, 24.1% in France) and "Other parents I know are choosing not to vaccinate" (31.9% in the US, 32.1% in France).

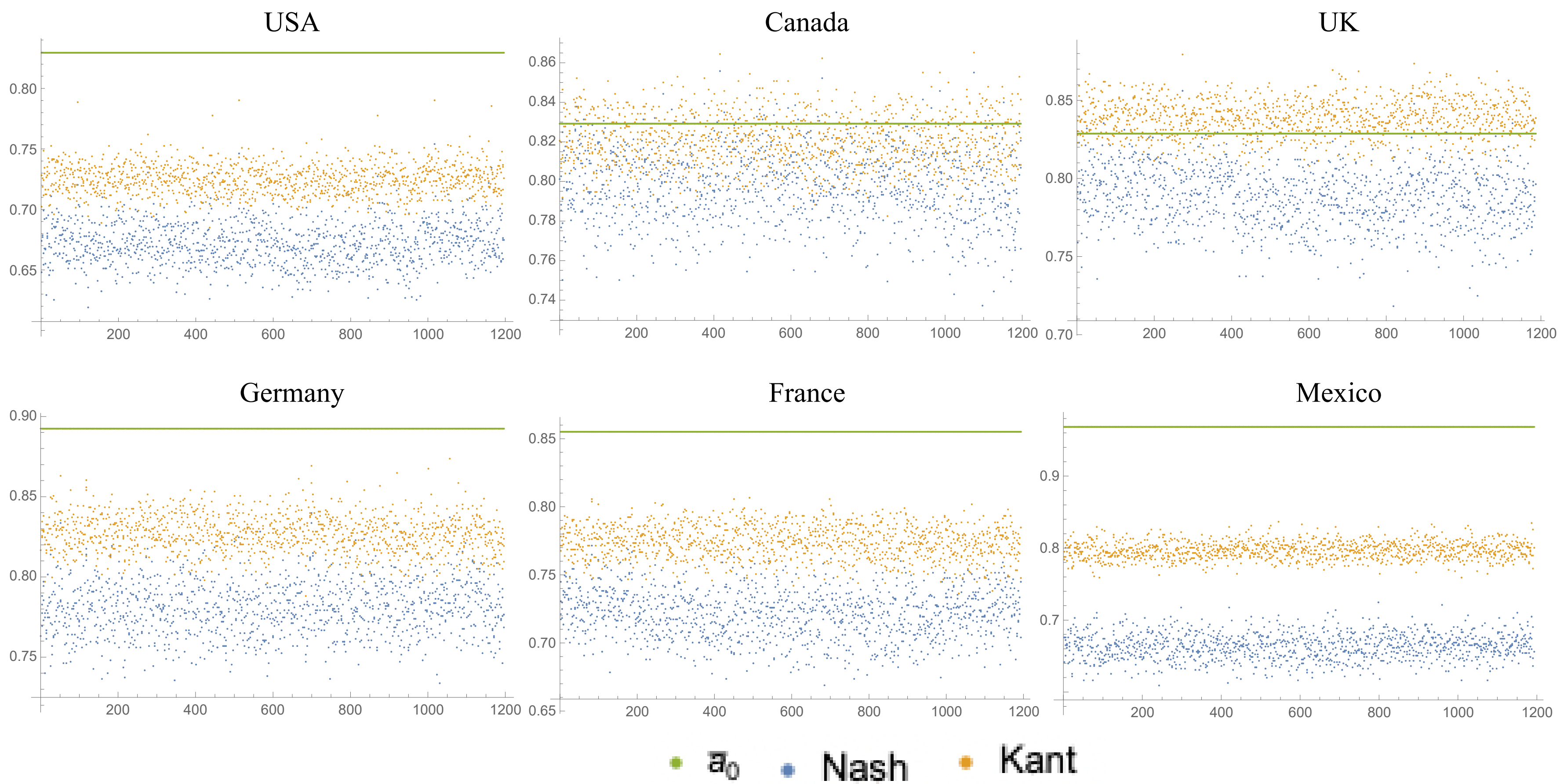

Figure 8. Coverage rates at the survey sample ($\bar{a}_0$), at the Nash equilibrium strategies ($\bar{a}^N(q^{*N})$), and at the Kantian equilibrium strategies ($\bar{a}^K(q^{*K})$). For all countries, $\bar{a}^N(q^{*N}) < \bar{a}^K(q^{*K})$. For all countries, except for Canada and the UK, $\bar{a}^N(q^{*N}) < \bar{a}^K(q^{*K}) < \bar{a}_0$. See the Online Appendix for numerical comparisons.

7 Conclusion

The vaccination of children can be modelled as a game in which there are significant positive externalities from the individual's choice to vaccinate; we say the game is monotone increasing. The Nash equilibria of such games are inefficient, a fact colloquially known as the free-rider problem. The Kantian equilibria of such games are generically efficient. We have shown, in a sample of six countries, that the Kantian model provides a superior explanation of vaccination behavior compared to the Nash model.

This might seem surprising: can we argue that some social institution is guiding parents to optimize in the Kantian manner? We do not do so here. However, we do conjecture that a fairly strong social norm is operative in many countries that induces parents to vaccinate at higher rates than they would were they playing the Nash equilibrium of the game. The social norm shares this property with the Kantian equilibrium, which, we argue, is the reason the Kantian model is more predictive of real behavior than the Nash model. Indeed, the Kantian equilibrium may be viewed as one precise formulation of a social norm.

The same reasoning applies to monotone *decreasing* games –ones in which each player's contribution imposes a negative externality on the welfare of others. The Nash equilibrium in monotone decreasing games is also inefficient, a fact known colloquially as the tragedy of the commons. The canonical fishing game, in which an increase in each fisher's labor decreases the productivity of the fishery for other fishers, is the textbook example. Elinor Ostrom (1990) studied common-pool resource problems in many communities around the world, and found that the participants often achieve efficient solutions, in the face of Nash equilibria that suffer from the tragedy of the commons, through the evolution of social norms that induce them to internalize the externalities of their behavior. Unfortunately, when Ostrom worked, the Kantian tool was not yet available: we suggest its fruitful application to the problems she studied.

APPENDIX A. The inefficiency of Nash equilibrium in strictly monotone games

Proposition 1 *Let $V^i : I^1 \times I^2 \times \ldots \times I^n \to \Re$ be differentiable payoff functions for $i = 1, 2, \ldots, n$ for an n-player strictly monotone game, where I^i is a non-negative real interval. Then any interior Nash equilibrium of the game is Pareto inefficient.*

Proof:

1. The conditions for Pareto efficiency of an interior Nash equilibrium $(x^1, \ldots, x^n) \in \Re^n_{++}$ are given by the solution of the following program:

$$\begin{gathered} \underset{(x^1, x^2, \ldots, x^n)}{Max} V^1(x^1, \ldots, x^n) \\ subject\ to \\ (\forall j = 2, \ldots, n)\big(V^j(x^1, \ldots, x^n) \geq k^j\big) \qquad (\lambda^j) \end{gathered} \tag{A.1}$$

2. The Kuhn-Tucker conditions for the solution of (A.*1*) are:

$$\begin{aligned} &V_1^1 + \lambda^2 V_1^2 + \cdots + \lambda^n V_1^n = 0 \\ &V_2^1 + \lambda^2 V_2^2 + \cdots + \lambda^n V_2^n = 0 \\ &\ldots \\ &V_n^1 + \lambda^2 V_n^2 + \cdots + \lambda^n V_n^n = 0 \end{aligned} \tag{A.2}$$

where $V_j^i = \frac{\partial V^i}{\partial x^j}$ for all i, j.

3. Suppose that $n \geq 3$. Assume that the game is strictly monotone increasing. By the interiority of the equilibrium, we have $V_i^i(x) = 0$ for all $i = 1, \ldots, n$. By monotonicity of the game, $V_j^i > 0$ for all (i, j) with $j \neq i$. Hence, we can rewrite the first two equations in

$$\begin{aligned} &V_1^1 + \lambda^2 V_1^2 + \cdots + \lambda^n V_1^n = 0 \\ &V_2^1 + \lambda^2 V_2^2 + \cdots + \lambda^n V_2^n = 0 \\ &\ldots \\ &V_n^1 + \lambda^2 V_n^2 + \cdots + \lambda^n V_n^n = 0 \end{aligned}$$

(A.2) as:

$$\begin{aligned} \lambda^2 V_1^2 + \lambda^3 V_1^3 + \cdots + \lambda^n V_1^n = 0 \ \ (since\ V_1^1 = 0) \\ \lambda^3 V_2^3 + \cdots + \lambda^n V_2^n = -V_2^1 \ \ (since\ V_2^2 = 0) \end{aligned} \tag{A.3}$$

By the positivity of V_1^j and the non-negativity of λ^j for all $j > 1$, we immediately have from the first equation in λ2V12+λ3V13+…+λnV1n= 0 $(since\ V_1^1 = 0)$⍰ $\lambda^3 V_2^3 + \cdots + \lambda^n V_2^n = -V_2^1 \ (since\ V_2^2 = 0)$⍰ (A.*3*) that $\lambda^j = 0$ for all $j = 2, \ldots, n$.

Therefore the second equation in λ2V12+λ3V13+…+λnV1n= 0 $(since\ V_1^1 =$

0) $\lambda^3 V_2^3 + \cdots + \lambda^n V_2^n = -V_2^1$ $(since\ V_2^2 = 0)$ (A.3) says $0 = -V_2^1$, a contradiction to strict monotonicity that establishes the result.

4. The case of $n = 2$ is disposed of even more quickly. The case of strictly monotone decreasing games has the same proof with a change of sign. n

APPENDIX B: Proof of Proposition 3

To prove the existence of such an equilibrium, we need to show that there is a number $\bar{a}^K$ such that the strategy profile defined by (2.10) indeed integrates to $\bar{a}^K$. In part A of the proof, we prove that if a value $\bar{a}^K$ exists which is consistent with this definition of the strategy profile, then eqn. (2.10) defines a Kantian equilibrium. In part B, we prove the existence of such a value of $\bar{a}^K$.

Part A. The strategy profile in (2.10) is a multiplicative Kantian equilibrium, if $\bar{a}^K$ exists consistent with this profile.

- Case 1 $\varepsilon < (1-u)\big(p - g(u,\bar{a})\big)$

(a) In this case, $\alpha^K(p,u,\varepsilon)$ is defined by the first branch of (2.10) Note that $\alpha^K \epsilon\ (0,1)$, since the probability of vaccinating on the first branch is strictly less than one.

(b) Calculate the derivative along the first branch of $\tilde{V}_{(p,u,\varepsilon)}$:

$$\frac{d^-}{d\rho}|_{\rho\nearrow 1}\tilde{V}^-_{(p,u)} = \alpha[p(u-1) + \varepsilon + p_0\pi(\bar{a}) + p_0\pi'(\bar{a})\bar{a}] - p_0\pi'(\bar{a})\bar{a}$$

(c) Calculate that on the interval $0 \le \rho \le \left(\frac{1}{\alpha^K(p,u,\varepsilon)}\right)$,

$$\frac{d^2\tilde{V}^-}{d\rho^2} = -(1-\rho\alpha^K)p_0\bar{a}^2\pi''(\rho\bar{a}) + 2\alpha^K p_0\pi'(\rho\bar{a})\bar{a},$$

which is negative on this interval, because by assumption π is a convex, decreasing function.

(d) Hence $\tilde{V}^-_{(p,u)}$ is a concave function of ρ on this interval. Observe that by definition of $\tilde{V}^-_{(p,u)}$ in (2.8), $\frac{d\tilde{V}^-_{(p,u)}}{d\rho} = 0$ at $\rho = 1$. Therefore the concave function $\tilde{V}^-_{(p,u)}$ is maximized for $\rho \in \left[0, \frac{1}{\alpha^K(p,u,\varepsilon)}\right]$ at $\rho = 1$.

(e) Next, we need to show that $\tilde{V}_{(p,u)}$ as a function of ρ is maximized at $\rho = 1$ on the interval $\rho \in \left[1, \frac{1}{\alpha^K(p,u,\varepsilon)}\right]$. This follows from part (d), because $\tilde{V}^+_{(p,u)}(\alpha, \bar{a}, \varepsilon; \rho)$ is dominated by $\tilde{V}^-_{(p,u)}(\alpha, \bar{a}, \varepsilon; \rho)$ on this interval (see (2.9)). We use the fact that the maximum of $\tilde{V}^-_{(p,u)}(\alpha, \bar{a}, \varepsilon; \rho)$ on the entire interval $\left[0, \frac{1}{\alpha^K(p,u,\varepsilon)}\right]$ is attained at $\rho = 1$. This establishes the claim for this case.

- Case 2 $\varepsilon \ge (1-u)\big(p - g(u,\bar{a})\big)$.

In this case, $\alpha^K = 1$. It is only necessary to maximize $\tilde{V}_{(p,u)}$ over the interval $\rho\epsilon[0,1]$, so we need only consult the left-hand derivative in part (b). Substituting $\alpha^K = 1$ into this expression, we have, in this case, that $\frac{d^-}{d\rho}|_{\rho\nearrow 1}\tilde{V}_{(p,u,\varepsilon)} \ge 0$. It follows immediately that $\tilde{V}$ is maximized at $\rho = 1$ in this case, and hence the proposed strategy profile is a Kantian equilibrium.

Part B. A value of $\bar{a}^K$ exists consistent with the strategy defined in (2.10).

The statistician sees only the average coverage rate for each tranche (p, u). This is given by integrating $\alpha^K dL(\varepsilon)$:

$$a^K(p,u) = \int_{-\infty}^{(1-u)\left(p-g(u,\bar{a}^K)\right)} \frac{-p_0\pi'(\bar{a}^K)\bar{a}^K}{(1-u)\left(p-g(u,\bar{a}^K)\right)-\varepsilon-p_0\pi'(\bar{a}^K)\bar{a}^K}\, dL(\varepsilon) +$$

$$1 - L\left((1-u)\left(p-g(u,\bar{a}^K)\right)\right). \qquad \text{(B.1)}$$

Then integrating over all (p, u):

$$\bar{a}^K = \iint_{-\infty}^{(1-u)\left(p-g(u,\bar{a}^K)\right)} \frac{-p_0\pi'(\bar{a}^K)\bar{a}^K}{(1-u)\left(p-g(u,\bar{a}^K)\right)-\varepsilon-p_0\pi'(\bar{a}^K)\bar{a}^K}\, dL(\varepsilon)dQ(p,u) +$$

$$\int \left[1 - L\left((1-u)\left(p-g(u,\bar{a}^K)\right)\right)\right] dQ(p,u), \qquad \text{(B.2)}$$

which is an equation in the single unknown $\bar{a}^K$. Existence requires showing that a value $\bar{a}^K$ exists satisfying (B.2). Define the expression on the right-hand side of (B.2) to be $z(\bar{a}^K)$. A fixed point of *z* is a solution of (B.2). Clearly the function *z* is continuous. We must show *z* maps the interval [0,1] into itself. The first (double) integral in the definition of *z* is less than $L\left((1-u)\left(p-g(u,\bar{a}^K)\right)\right)$, since the integrand is always less than one. Hence, by (B.2), the mapping *z* sends the unit interval *I* into itself. Since *z* is a continuous function, the Brouwer Fixed Point Theorem tells us that a solution $\bar{a}^K$ of (B.2) exists. □□

APPENDIX C: Estimation of the parameter γ

We use the source "WHO vaccine-preventable diseases: Monitoring system, 2020 global summary," https://apps.who.int/immunization_monitoring/globalsummary/, which contains data for a large set of countries on infectious disease immunization rates and morbidity.

A *cohort* of children is the set of children in the country born in a given year.

For a particular country, let:

$n' =$ total population of children ages 0-5 in year $t, t = 2015, \dots, 2019$

$r' =$ measles immunization coverage rate, children under 5, year *t*

$c' =$ number of measles cases, year *t*

$\bar{u} =$ number of susceptible children under 5 in a given cohort

$\bar{n} = \frac{\sum_{t=1}^{5} n^t}{5}$; $\bar{n}/_{5} =$ number of children in a given cohort

$\bar{r} = \frac{\sum_{t=1}^{5} r^t}{5}$

$\bar{c} = \frac{\sum_{t=1}^{5} c^t}{5}$

$p =$ probability that a susceptible child of a given cohort contracts measles in a given year

$\pi =$ probability that a susceptible child of a given cohort contracts measles by five years of age

By definition, $\bar{u} = \frac{\bar{n}}{5}(1 - \bar{r})$. The median age of contracting measles is age five. Therefore, the number of cases of measles of children under five in a given cohort in a given year is $\frac{\bar{c}}{10}$. Therefore $p = \frac{\bar{c}/10}{\bar{u}} = \frac{\bar{c}}{10\bar{u}}$.

Assume that an unvaccinated (susceptible) child in a given cohort has a probability *p* of contracting measles in each year under five. Then:

$$\pi = p + p(1-p) + \cdots + p(1-p)^4 = p\frac{1-(1-p)^5}{p} = 1 - (1-p)^5.$$

In our model we have $\pi(r) = (1-r)^{\gamma}$. We propose that $\pi(r)$ is precisely the value π defined above: as a parent, I am concerned with the probability that my young child contracts measles if I choose not to vaccinate her, knowing that the coverage rate is *r*.

As described in the text, we assume the contagion process in North America (Canada and the US) is different from in Europe (UK, Germany, France). For each country j, we compute a data point (r, π). Hence, we compute two values of γ: γ^{NA} gives the best fit of the function $\pi(\cdot)$ to the points $\{(r^{\mathrm{US}}, \pi^{\mathrm{US}}), (r^{\mathrm{Can}}, \pi^{\mathrm{Can}})\}$ and γ^{EUR} gives the best fit of the function $\pi(\cdot)$ to the points $\{(r^j, \pi^j) | j \in \{\mathrm{UK}, \mathrm{France}, \mathrm{Germany}\}\}$. See Figure B.1 and Figure B.2.

Unfortunately, the data set does not provide measles morbidity for Mexico.

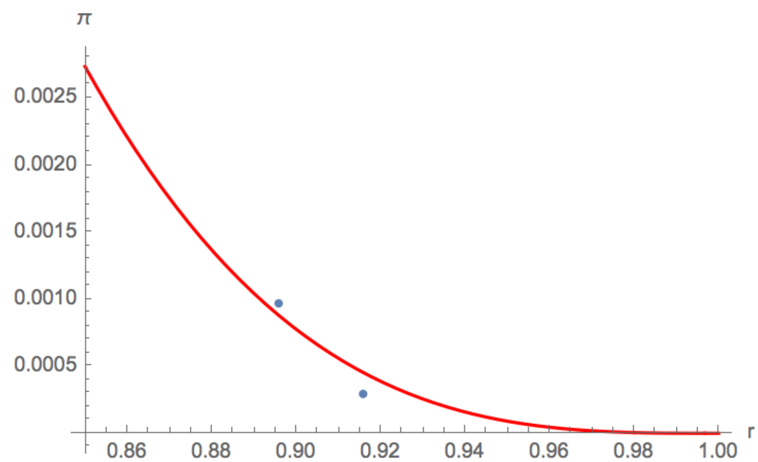

Figure B.1 Fitting the function $\pi(\cdot)$ for the US and Canada: $\gamma^{NA} = 3.110$.

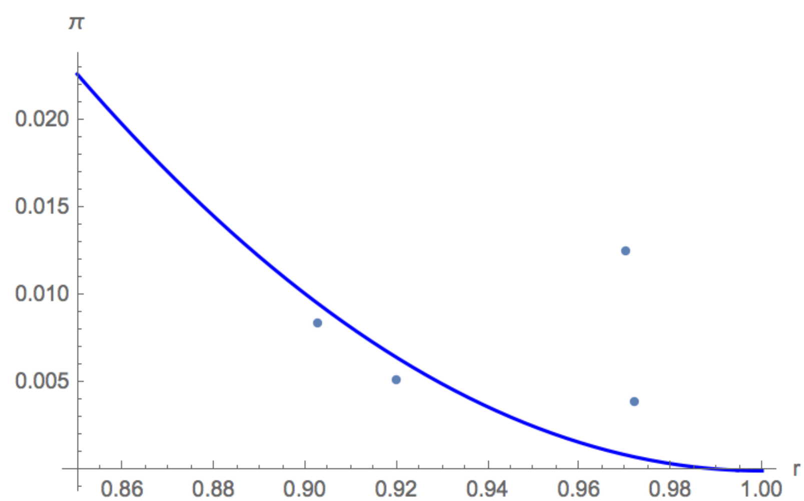

Figure B.2 Fitting the function $\pi(\cdot)$ for the four European countries (Sweden included): $\gamma^{EUR} = 1.995$.

References

Alm, J. 2019. "What motivates tax compliance?" *Journal of Economic Surveys* 33, 353-388.

Auld, C. 2003. "Choices, beliefs, and infectious disease dynamics" *Journal of health economics* 22.3, 361-377.

Avery, C., W. Bossert, A. Clark, G. Ellison and S.F. Ellison, 2020. "An economist's guide to epidemiology models of infectious disease," *Journal of Economic Perspectives.* 34, 79-104.

Bauch, C., A. Galvani and D. Earn, 2003, "Group interest versus self-interest in smallpox vaccination policy" *Proceedings of the National Academy of Sciences* 100.18, 10564-10567.

Bauch, C., and D. Earn, 2004, "Vaccination and the theory of games" *Proceedings of the National Academy of Sciences* 101.36, 13391-13394.

Brito, D., E. Sheshinski, and M. Intriligator. 1991, "Externalities and compulsary vaccinations", *Journal of Public Economics* 45.1, 69-90.

Crainich, D., L. Eeckhoudt, and M. Menegatti. 2019, "Vaccination as a trade-off between risks", *Italian Economic Journal* 5.3, 455-472.

Dabbagh, A., R.L. Laws et alii. 2018. "Progress toward regional measles elimination –Worldwide, 2000-2017," *Morbidity and Mortality Weekly Report* 67 (47), US Dept. of Health and Human Services, Center for Disease Control and Prevention.

d'Albis H., J. Etner and J. Thuilliez, 2022, "Vaccination under pessimistic expectations in clinical trials and immunization campaigns", mimeo.

Fine, P., and J. Clarkson, 1986, "Individual versus public priorities in the determination of optimal vaccination policies", Am. J. Epidemiol. 124, 1012–1020

Gatwood, J., McKnight, M., Fiscus, M., Hohmeier, K. C., & Chisholm-Burns, M. (2021). Factors influencing likelihood of COVID-19 vaccination: A survey of Tennessee adults. *American Journal of Health-System Pharmacy*, *78*(10), 879-889.

Geoffard, P-Y, and T Philipson, 1997, "Disease eradication: private versus public vaccination", *The American Economic Review* 87.1, 222-230.

Holt, C.A. and S.K. Laury, 2002. "Risk aversion and incentive effects," *American Economic Review* 92, 1644-1655.

Kan, T., & Zhang, J. (2018). Factors influencing seasonal influenza vaccination behaviour among elderly people: a systematic review. *Public health*, *156*, 67-78.

Kreps, S., Prasad, S., Brownstein, J. S., Hswen, Y., Garibaldi, B. T., Zhang, B., & Kriner, D. L. (2020). Factors associated with US adults' likelihood of accepting COVID-19 vaccination. *JAMA network open*, *3*(10), e2025594-e2025594.

Nuscheler, R. and K. Roeder, 2016, "To Vaccinate or to Procrastinate? That is the prevention question", *Health economics* 25.12, 1560-1581.

Ostrom, E. 1990. *Governing the commons: The evolution of institutions for collective action*, Cambridge, UK: Cambridge University Press.

Philipson, T., 2000, "Economic epidemiology and infectious diseases", *Handbook of health economics* 1, 1761-1799.

Roemer, J.E. 2019. *How we cooperate: A theory of Kantian optimization,* New Haven: Yale University Press.

Shim, E., G. Chapman, J. Townsend and A. Galvani, 2012a, "The influence of altruism on influenza vaccination decisions", *Journal of The Royal Society Interface* 9.74, 2234-2243.

Shim, E., Grefenstette, J., S. Albert, B. Cakouros and D. Burke, 2012b, "A game dynamic model for vaccine skeptics and vaccine believers: measles as an example", J. Theor. Biol. 295, 194–203.

Strebel, P. M. and W. A. Orenstein. 2019. "Measles," *New England Journal of Medicine.* 381(4), 349–357. PMID: 31184814.

Vanlandingham, M. J., Suprasert, S., Grandjean, N., & Sittitrai, W. (1995). Two views of risky sexual practices among northern Thai males: The health belief model and the theory of reasoned action. *Journal of Health and Social behavior*, 195-212.

Yang, Z. J. (2015). Predicting young adults' intentions to get the H1N1 vaccine: An integrated model. *Journal of Health Communication*, *20*(1), 69-79.