Hypocrisy and Hypocrites: A Game-Theoretic Note

Bruce Wydick
University of San Francisco

Abstract: Hypocrisy is the feigning of beliefs or virtues that one does not truly possess. Hypocrisy among people claiming religious faith is often provided as a justification for non-religious people to eschew religion. The claim is that widespread hypocrisy among people claiming religious faith is evidence that such beliefs are either false or have no impact on behavior. This short paper provides a proof of precisely the opposite: the existence of genuine faith is a necessary condition for hypocrisy. It demonstrates that the existence of hypocrisy is evidence that genuine religious faith has produced salient differences between non-believers and genuine believers. This is illustrated in a game-theoretic model of market exchange when agent types are hidden. In the resulting Nash equilibrium, hypocrisy always lingers in the shadow of genuine faith when costs of religious profession are low.

“Tis curious that we only believe as deeply as we live.” Ralph Waldo Emerson

“Of all bad men religious bad men are the worst.” C.S. Lewis

“The hypocrite, certainly, is a secret atheist; for if he did believe there was a God, he durst not be so bold as to deceive Him to His face.”

Thomas Adams

Author’s Note: I thank the editors of Faith & Economics and two anonymous referees for comments that substantially improved the manuscript. Department of Economics, University of San Francisco, 2130 Fulton Street, San Francisco, CA 94117; Visiting Professor, Department of Agricultural and Resource Economics, University of California at Berkeley; e-mail: wydick@usfca.edu.

© 2012 Association of Christian Economists
Recent books by outspoken atheists have pointed to the existence of hypocrisy among religious people as evidence against the veracity and genuineness of religious faith in general. Christopher Hitchens’ *god is Not Great* (2007) devotes an entire chapter to supporting the hypothesis that hypocrisy among people religious people provides evidence that religious people behave no better than non-religious people. Similarly Richard Dawkins’ *The God Delusion* (2006) maintains that evil acts carried out by people claiming religion constitute evidence that religious faith fails to have a positive effect on human behavior.

This paper is a response to such arguments. In a game-theoretic model, I show that the existence of hypocrisy does not constitute valid evidence against the veracity or genuineness of religious faith. Indeed, contrary to the arguments of recent commentators on religion, the case is actually the opposite: religious hypocrisy can rationally exist only in the presence of genuine faith, in which the actions of believers are on average viewed more positively than those of non-believers. When claims of religious faith are relatively costless to individuals, hypocrisy will invade genuine belief, resulting in a Nash outcome in which equilibrium behavior between the two groups may be virtually identical. This result also helps to explain empirical data such as McEleny and Barro (2006) and other survey results (such as Barna 2002; Barna Group 2004; Sider 2005) that appear to find little difference in behavioral outcomes between believers and non-believers.

The model I present is set in the framework of exchange under imperfect information, a context that provides scope for cheating. While the context in which I present the model is market exchange, the model is generalizable to broader contexts such as social exchanges within a friendship, courting and marriage, and other types of non-economic interaction. From the Nash equilibrium condition derived in the model, we can derive four major points about hypocrisy. First, genuine religious belief that positively affects behavior in a manner that is salient to outside agents is a necessary condition for hypocrisy. Second, the greater the difference in average behavior between genuine believers and non-believers, the greater will be the prevalence of hypocrisy. Third, hypocrisy will be abundant in societies and contexts with greater religious freedom; its presence diminishes to zero as...
religious persecution increases. Lastly, the model shows that when genuine belief is unobservable, it is impossible for reduced-form observational studies to estimate causal effects of religion on behavior.

A Model of Hypocrisy

Consider an economy of risk-neutral agents indexed by \( i \in \{1 \ldots n\} \) who engage in one-shot transactions with one another under both uncertainty and imperfect information. An agent may be either a buying agent or a selling agent in any particular exchange. There is one-way moral hazard in the exchange that puts a buying agent at risk. A transaction yields 1 unit of utility to a buying agent if the selling agent complies in the transaction, but zero units of utility if the selling agent fails to comply. A selling agent may fail to comply in a transaction either because he is dishonest, in which case he receives some measure of added utility from non-compliance, or due to circumstances beyond his control. Examples of such transactions might be a credit transaction in which a builder skims on construction materials, sells a buyer shoddy goods, or agreements to produce a good or service by a future date in which the service is rendered late or not at all.

There are two types of agents, “Believers” (\( B \)), who believe God will ultimately hold them accountable for moral decisions, and “Non-believers” (\( N \)) who do not believe this. Precisely how \( B \) agents believe God will hold them accountable, whether in this life or in the afterlife, is tangential to the model. What is important is that Believers believe the repercussions for dishonesty outweigh the gain in utility from cheating.

In the population of agents, \( \mathcal{A} \), we will assume the existence of a subset of \( n_B \) believers, \( B \subset \mathcal{A} \) and \( n_N \) non-believers, \( N \subset \mathcal{A} \), where \( B \cup N = \mathcal{A} \). Let \( \gamma_i \) be the exogenous probability that a (selling) agent \( i \) complies in a transaction. Assume that the probability distribution of compliance for \( B \), \( g_B(\gamma_i) \) is higher than for \( N \), \( g_N(\gamma_i) \), such that for their respective distributions functions, \( G_N(\gamma_i) > G_B(\gamma_i) \) for \( \forall i \). This does not assume that every \( B \) agent has a higher probability of complying than every \( N \) agent, but that \( B \) agents display a higher rate of compliance on average than \( N \) agents.

A buying agent discounts the price that she is willing to pay for the one-unit exchange by the probability that the selling agent will not comply in the transaction. Thus the competitive price that any buying agent \( i \) is willing to pay in an exchange with a selling agent \( j \) is equal to
the expected rate of compliance of the selling agent, or \( p_j = E(y_j) \). In a situation where there is full information, each type with \( \mathcal{A} \) receives a price \( p_j \) associated with his type.

But now assume that there is asymmetric information over types. Prior to the transaction, all selling agents can signal whether they are \( B \) or \( N \). Because \( G_N(y_i) > G_B(y_i) \), all agents would like to be able to signal that they are \( B \) agents. We will assume that for \( B \) agents, the cost of this religious signal, \( r_i = 0 \). For \( N \) agents, the cost of the signal is \( r_i = c \), where we will assume \( c \in (0, E(y_B) - E(y_N)) \) such that an interior solution exists.

Examples of the religious signal in, for example, a Christian context, would be church attendance and attendance at prayer meetings and Bible studies, baptism, using appropriately religious language, public professions of faith, and so forth. These activities are costless to the believer in the sense that they would be undertaken for their own sake as a natural consequence of belief, where this belief is accompanied by a perception of divine accountability that steers the believer to honesty. A person who engages in these activities without genuine belief is unlikely to consciously view himself as a hypocrite. Rather he will have observed that “these are things that ‘good people’ do,” and, wanting to be associated with ‘good people,’ will imitate these behaviors himself, but at some cost: He will find these activities generally uninteresting, yet as a kind of sacrifice necessary to maintaining a positive public impression. However, while in religious settings this person openly professes a level of personal integrity associated with genuine belief, in private dealings his behavior reflects that of one who holds no regard for the ultimate accountability of his actions; he is a hypocrite.

The first thing to notice is that provided \( c < E(y_B) - E(y_N) \), not all agents sending the religious signal will be members of \( B \). Some number \( n_{NB} \) of \( N \) will falsely send the religious signal. The equilibrium condition will be that

\[
\frac{n_B \int_0^1 y_B g_B(y) dy_B + n_{NB} \int_0^1 y_N g_N(y) dy_N}{n_B + n_{NB}} - c = \int_0^1 y_N g_N(y_i) dy_N .
\]

The explanation for the equilibrium is the following: Observe that in the equilibrium in (1) there will be a flow of \( N \) types into \( NB \) so that \( n_{NB} \) will grow (representing an increase in hypocrisy) until the expected gain in transactions by the \( N \) types is the same from either being religious or
non-religious. In other words, as long as there exist gains to being religious, more and more \( N \) types will become falsely ‘religious’ (\( NB \) types) until the behavior of those associated with religion is sufficiently diluted, such that in equilibrium the behavior of those associated with religion and those not associated with religion is nearly the same, where any difference in behavior between the two groups a positive function of \( c \). Re-arranging (1), substituting mean levels of compliance for each group, algebraically re-arranging, and then substituting in the fractions \( f_B, f_{NB} \) of those claiming religious faith, we can re-write (1) as

\[
f_{NB} = \frac{f_B}{c} (\bar{y}_B - \bar{y}_N - c)
\]

Equation (2) yields the following proposition:

**Proposition:** Hypocrisy can only exist in the presence of genuine faith.

The proof to the proposition is clear from (1) and (2). If \( \bar{y}_B - \bar{y}_N \leq c \), then \( f_{NB} \), the fraction of Non-Believers claiming religious faith, is zero. The only rational motivation for spurious imitation is that there exists something genuine to imitate. Without the existence of genuine faith and a behavioral premium among those with such a faith, hypocrisy cannot rationally exist.

**Corollary 1:** Hypocrisy is more prevalent when the behavior of Believers differs more substantially from the behavior of Non-believers.

This is also clear from (2), where \( \frac{\partial f_{NB}}{\partial \bar{y}_B} > 0 \). As \( \bar{y}_B - \bar{y}_N \) becomes greater, the Nash equilibrium described in (1) and (2) is characterized by a greater fraction of hypocrites in the population among those who claim religious faith.

**Corollary 2:** Hypocrisy is less prevalent when it is more costly for non-believers to claim religion.

Consider the difference between an atmosphere of persecution and one of religious freedom. This increases \( c \) for \( N \) agents. Since as seen in (2), \( \frac{\partial f_{NB}}{\partial c} < 0 \), with greater persecution of religious people, there is less hypocrisy in the resulting equilibrium. Hypocrisy should be less abundant in countries where there is strong persecution of religion.
COROLLARY 3: The equilibrium behavior of those claiming religious belief will not differ significantly than the behavior of others.

A number of studies have reported “surprising” findings that the behavior of people who claim religious faith differs little from those who do not purport to have faith (see Sider 2005 for an excellent review). The implication of some from these empirical studies is that religious faith has no discernible effect on behavior. What is clear here is that observationally similar behavior between those claiming religious faith and others who do not has no implication for the causal effect of genuine belief on behavior. Because what we observe is an equilibrium condition, we cannot make statistical inferences on the causal effect of religion on behavior using observational data in reduced-form estimations. Moreover, if genuine faith is unobservable, it is impossible to empirically measure the causal effect of faith observationally from behavioral outcomes without a structural instrument that is correlated with genuine belief but otherwise uncorrelated with behavioral outcomes.

Concluding Comments

In the introduction to Jonathan Malesic’s (2009) Secret Faith in the Public Square: An Argument for the Concealment of Christian Identity, he relates the story of a contractor submitting a bid on his in-law’s house, in which the contractor made an effort to point out that he was a member of the First Baptist Church in a nearby town. While Malesic argues that sociologists such as Weber have noted “religious identity has long been a form of currency in American society” (p.13), his book highlights the negative role that public professions of faith have had on the perception of the church, calling for Christians to practice a silent faith in order to restore the credibility of the church in contemporary society. This incentive to signal honesty in a context where there is scope for moral hazard, and where religion is associated with honest behavior, forms the context of this model.

I point out that hypocrisy is unlikely to constitute a conscious behavior. In most instances it may function as a type of social norm, the product of individuals desiring to be regarded as “good people” even in the absence of genuine belief. In this way, the behavioral patterns of hypocrisy are likely to develop subconsciously and may not involve the rational calculus assumed in most economic models. For example, an equilibrium such as in (1) and (2) can result when those claiming
religion are a mixture of genuine believers and “traditional” adherents, where in practice the behavior of the latter conforms closely to that of non-religious people. These “traditional adherents” may be quick to disassociate themselves from the religion if the average behavior among those claiming religion was viewed to be worse than that among those claiming no religious belief. But provided the cost of the religious signal remains below the difference in behavior between those claiming and not claiming religious belief, it remains advantageous to claim religion. Thus an out-of-equilibrium outcome in which a subset of individuals can be identified as more trustworthy, and benefit from this perceived trustworthiness, is unlikely to be sustainable in contexts where religious talk is cheap. In such contexts individuals may unconsciously support a Nash equilibrium characterized by widespread nominal belief where hypocrisy is prevalent. It is this kind of equilibrium that troubles theologians such as Malesic, but that it is possible to eradicate hypocrisy in a context where there is religious freedom seems unlikely: the more substantial the behavioral differences between believers and unbelievers, the greater the incentives are for hypocrisy to emerge.

The widespread existence of hypocrisy among people claiming religion does not disprove the behavioral effects stemming from genuine faith. On the contrary, hypocrisy proves its existence because genuine, behavior-changing faith is a necessary condition for hypocrisy to exist. It is twenty dollar bills that are counterfeited, not pennies.

References


