



DEPARTAMENTO DE CIÊNCIAS ECONÔMICAS

TEXTO PARA DISCUSSÃO Nº 12

**PATRICK JANE NÃO LÊ MENTES, FAZ SCREENING:
THE MENTALIST E A ECONOMIA DA INFORMAÇÃO**

Otávio Florentino Detoni

Agosto de 2026

Bibliotecária responsável: Celsimara M. do C. Guimarães (CRB 6: 1963)

Setor de Processamento Técnico da Divisão de Bibliotecas da Universidade Federal de São João del-Rei

Detoni, Otávio Florentino

D482p Patrick Jane não lê mentes, faz screening: The Mentalist e a economia da informação [recurso eletrônico] / Otávio Florentino Detoni . – São João del-Rei, MG: UFSJ / Departamento de Ciências Econômicas, 2026.

1 recurso on-line (15 p.) . (Texto para discussão ; nº 12).

Acesso em: https://ufsj.edu.br/dceco/textos_para_discussao.php

Inclui referências.

1. Informação, teoria da (Economia). 2. Economia – Estudo e ensino. I. Título.

CDU 330

UNIVERSIDADE FEDERAL DE SÃO JOÃO DEL-REI
DEPARTAMENTO DE CIÊNCIAS ECONÔMICAS

Patrick Jane não lê mentes, faz *screening*

The Mentalist e a economia da informação

Otávio Florentino Detoni

Professor do Departamento de Ciências Econômicas (DCECO/UFSJ)

otaviodetoni@ufs.br

*Texto de economia da informação preparado para uso didático
e disponibilizado para discussão acadêmica.*

DCECO/UFSJ
São João del-Rei
2026

Sumário

1. O truque de Jane: criar uma escolha	2
2. Como aprender sobre alguém que sabe mais do que você?	2
2.1. Observar características	3
2.2. Ouvir o que a pessoa diz	3
2.3. Provocar uma escolha	3
3. Antes do modelo: como atualizar uma suspeita	3
3.1. Um exemplo com cem suspeitos	4
3.2. A mesma conta pela regra de Bayes	4
4. O que é <i>screening</i> ? O exemplo do seguro	5
4.1. O problema da seguradora	5
4.2. A solução: oferecer duas opções	5
4.3. As três condições do mecanismo	5
4.4. A informação tem custo	6
4.5. Sinalização não é <i>screening</i>	6
5. O anúncio de Jane como <i>screening</i>	6
6. O modelo básico: ficar parado ou reagir	7
6.1. O custo de cada ação	7
6.2. Derivando a condição de reação	7
6.3. Exemplo numérico	8
6.4. Agir não é o mesmo que ser observado	8
6.5. O paradoxo da vigilância perfeita	8
7. Uma reação não prova culpa	9
8. Credibilidade: o blefe de hoje enfraquece o de amanhã	9
9. Por que Red John é um adversário diferente?	10
9.1. Menor crença no anúncio e maior vigilância	10
9.2. Quando exposição também é recompensa	10
9.3. Bayes e a importância do ponto de partida	11
10. O mesmo mecanismo na economia real	11
10.1. Fiscalização tributária, trabalhista ou ambiental	11
10.2. Auditoria e <i>compliance</i>	12
10.3. Diplomas, selos e reputação em plataformas	12
10.4. Exposição pública e incentivos perversos	12
11. Síntese do argumento	13
12. Conclusão	13
Referências	15

RESUMO

Em *The Mentalist*, Patrick Jane muitas vezes anuncia que uma prova está prestes a aparecer e observa quem tenta impedir a investigação. Este texto usa essas cenas para examinar um problema central da economia da informação: como aprender sobre alguém quando sua característica relevante não pode ser observada diretamente. O argumento é construído passo a passo. Primeiro, distinguem-se pistas, declarações e escolhas induzidas. Em seguida, o conceito de *screening* é apresentado com o exemplo do mercado de seguros. Por fim, o anúncio de Jane é transformado em um modelo simples, no qual o suspeito escolhe entre ficar parado e reagir. O modelo mostra quando a armadilha separa tipos, por que uma reação não prova culpa, como o uso repetido de blefes reduz sua credibilidade e por que um adversário como Red John enfrenta incentivos diferentes. Exemplos numéricos e aplicações à fiscalização, à auditoria, às certificações e às plataformas tornam a exposição acessível sem abandonar o argumento formal.

Palavras-chave: economia da informação; *screening*; ensino de economia.

Classificação JEL: A22; D82; D83; K42.

ABSTRACT

In *The Mentalist*, Patrick Jane often announces that evidence is about to emerge and watches who tries to obstruct the investigation. This paper uses those scenes to examine a central problem in information economics: how to learn about someone when the relevant characteristic cannot be observed directly. The argument is developed step by step. It first distinguishes clues, statements, and induced choices. It then introduces screening through the standard insurance-market example. Finally, Jane's announcement is translated into a simple model in which a suspect chooses between remaining still and reacting. The model explains when the trap separates types, why a reaction does not prove guilt, how repeated bluffs erode credibility, and why an informed adversary such as Red John faces different incentives. Numerical examples and applications to enforcement, auditing, certification, and platforms make the exposition accessible without abandoning the formal argument.

Keywords: economics of information; screening; economics education.

JEL Classification: A22; D82; D83; K42.

1. O truque de Jane: criar uma escolha

Uma cena se repete em *The Mentalist*. Patrick Jane diz aos suspeitos que a polícia encontrou uma prova decisiva ou que está prestes a realizar um teste capaz de identificar o culpado. O anúncio nem sempre é verdadeiro. Mesmo assim, alguém tenta destruir um objeto, avisar um cúmplice, alterar uma história ou fugir.

À primeira vista, Jane parece ter descoberto a mentira por meio de uma expressão facial ou de um gesto nervoso. Essa leitura perde a parte mais interessante do método. Em muitas cenas, ele não se limita a observar uma reação que surgiria de qualquer maneira. Ele cria uma situação nova e faz com que cada pessoa precise escolher o que fazer.

Antes do anúncio, culpados e inocentes podem repetir a mesma frase: “eu não fiz nada”. Depois do anúncio, as consequências de agir não são iguais para todos. Quem acredita que perderá muito se a prova aparecer pode tentar impedir a investigação. Quem não teme a descoberta tem menos razão para assumir o custo e o risco de reagir.

Essa é a ideia central do texto. Patrick Jane não precisa observar diretamente quem é culpado. Ele muda os incentivos e cria uma escolha. Como pessoas com informações e interesses diferentes podem escolher ações diferentes, a própria escolha produz informação.

O mesmo problema aparece fora da ficção. Uma seguradora não observa perfeitamente o risco de cada motorista. Uma empresa não conhece toda a qualidade de um candidato. Uma plataforma não sabe de antemão quais vendedores são confiáveis. Uma autoridade fiscal não observa todas as irregularidades. Em todos esses casos, o lado menos informado procura criar regras, contratos ou testes que façam os diferentes tipos se comportarem de maneira distinta.

A escolha de *The Mentalist* também tem uma motivação pessoal. Foi a série que me introduziu ao hábito de acompanhar seriados de televisão e permanece entre as minhas preferidas. Exibida originalmente pela CBS entre 2008 e 2015, a produção voltou a alcançar amplo público quando suas sete temporadas chegaram ao catálogo da Netflix em 2026.

Há uma ironia útil nessa retomada. Uma série sobre um investigador que produz informação ao observar escolhas voltou ao primeiro plano por meio de uma plataforma que também tenta aprender sobre preferências a partir das escolhas de seus usuários.

O uso de filmes e séries no ensino de economia possui literatura própria (Becker e Watts 1996; Leet e Houser 2003; Watts e Becker 2008; Luccasen e Thomas 2010; Ghent, Grant e Lesica 2011). Aqui, porém, a cena não serve apenas para tornar a aula mais atraente. Ela reproduz a estrutura do problema econômico. O modelo formal ajudará a mostrar o que a intuição do espectador acerta e, principalmente, onde ela pode enganar.

2. Como aprender sobre alguém que sabe mais do que você?

Existe *informação assimétrica* quando uma parte conhece algo relevante que a outra não conhece. O suspeito sabe se cometeu o crime; Jane não. O motorista conhece melhor seu modo de dirigir; a seguradora não. O vendedor sabe mais sobre a qualidade do produto; o comprador não.

Vamos chamar a característica escondida de *tipo*. No exemplo mais simples, o suspeito pode ser de dois tipos:

$$\theta \in \{C, I\},$$

em que C representa o culpado e I representa o inocente. O tipo é informação privada da própria pessoa.

O lado desinformado pode tentar aprender de três maneiras: observando características, ouvindo declarações ou provocando escolhas. As três fontes de informação não são equivalentes.

2.1. Observar características

Jane presta atenção a calos, marcas, roupas, vocabulário e pequenos hábitos. Esses elementos podem revelar profissão, rotina ou relações pessoais. Eles são úteis porque não foram necessariamente produzidos com o objetivo de convencer o investigador.

Ainda assim, uma pista não determina o tipo. Nervosismo pode ser mais frequente entre culpados, mas um inocente também pode ficar nervoso durante um interrogatório. A pergunta correta não é “culpados ficam nervosos?”, e sim:

$$P(\text{nervosismo} \mid C) > P(\text{nervosismo} \mid I)?$$

Se a resposta for positiva, o nervosismo contém informação. Isso não significa que todo nervoso seja culpado. A pista altera uma probabilidade; não entrega uma certeza.

2.2. Ouvir o que a pessoa diz

Durante um interrogatório, quase todos têm interesse em dizer “sou inocente”. A frase custa pouco e pode ser imitada tanto por culpados quanto por inocentes. Na teoria dos jogos, uma mensagem desse tipo é chamada de *cheap talk*: uma fala barata, sem compromisso direto e sem custo relevante (Crawford e Sobel 1982; Farrell e Rabin 1996).

Uma declaração só informa quando não é igualmente fácil ou vantajoso para todos os tipos reproduzi-la. A antiga alegação de Jane de possuir poderes paranormais ilustra o problema. Qualquer pessoa podia fazer a mesma afirmação. Quando o sinal pode ser copiado sem dificuldade, seu valor informativo tende a desaparecer.

2.3. Provocar uma escolha

A terceira estratégia é diferente. O investigador altera o ambiente e observa como cada pessoa responde. Em vez de esperar que o tipo privado seja revelado espontaneamente, ele cria uma decisão cujas consequências variam entre os tipos.

Esse procedimento pertence à família do *screening*. Em português, o termo pode ser entendido como triagem ou seleção por autosseleção. O lado desinformado oferece opções ou cria uma regra; o lado informado escolhe; e a escolha revela alguma informação.

Compare as três situações

Fonte	Quem produz o indício?	Exemplo
Característica	A pessoa já traz o traço consigo	Calos nas mãos, uniforme, modo de falar
Declaração	O lado informado escolhe o que dizer	“Eu sou inocente”
Escolha induzida	O lado desinformado desenha a situação	Jane anuncia uma prova e observa quem reage

3. Antes do modelo: como atualizar uma suspeita

Suponha que alguém reaja ao anúncio de Jane. Quanto essa reação deve aumentar a crença de que a pessoa é culpada? A resposta depende de três informações:

1. a probabilidade inicial de culpa;
2. a frequência da reação entre culpados;
3. a frequência da reação entre inocentes.

É mais fácil começar com uma contagem do que com uma fórmula.

3.1. Um exemplo com cem suspeitos

Comecemos pelo caso concreto da série. Jane reúne quatro suspeitos, um dos quais cometeu o crime. Sem nenhuma outra informação, a suspeita inicial sobre cada um é de $1/4$, ou seja, 25%.

Contar culpados e inocentes em uma única investigação, porém, produziria frações de pessoa. Para enxergar as frequências com números inteiros, imagine cem investigações como essa. Em cada uma há quatro suspeitos e um culpado. Se acompanharmos um suspeito por investigação, teremos cem pessoas, das quais 25 são culpadas e 75 são inocentes, exatamente a proporção que o caso concreto estabelece.

Suponha agora que:

- 80% dos culpados reajam ao anúncio: $0,80 \times 25 = 20$ pessoas;
- 20% dos inocentes também reajam: $0,20 \times 75 = 15$ pessoas.

Ao todo, 35 pessoas reagem. Dessas, 20 são culpadas. Logo:

$$P(C | E) = \frac{20}{35} \approx 0,57,$$

em que E representa a reação observada.

A reação elevou a suspeita de 25% para aproximadamente 57%. Isso é muita informação, mas não é prova definitiva: quinze inocentes também reagiram.

Por ora, esses 20% são apenas um dado do exemplo. A seção 7 explica por que um inocente pode ter motivo para reagir.

3.2. A mesma conta pela regra de Bayes

Se $p = P(C)$ é a probabilidade inicial de culpa, a regra de Bayes escreve a conta anterior de forma geral:

$$\mu(C | E) = \frac{P(E | C) p}{P(E | C) p + P(E | I)(1 - p)}. \quad (1)$$

No exemplo:

$$\mu(C | E) = \frac{0,80 \times 0,25}{0,80 \times 0,25 + 0,20 \times 0,75} = \frac{0,20}{0,35} \approx 0,57.$$

O numerador mede a probabilidade de a pessoa ser culpada e reagir. O denominador mede a probabilidade total de uma reação, seja ela produzida por culpado ou inocente.

Convém distinguir duas perguntas parecidas. $P(E | C)$ pergunta: “se a pessoa é culpada, qual é a chance de ela reagir?” Já $P(C | E)$ pergunta: “se observei uma reação, qual é a chance de a pessoa ser culpada?” As duas probabilidades não são iguais. Confundir uma probabilidade condicional com sua inversa é conhecido na literatura jurídica como falácia do promotor, sobretudo quando uma baixa probabilidade de observar a evidência entre inocentes, $P(E | I)$, é apresentada como se implicasse uma baixa probabilidade de inocência diante da evidência, $P(I | E)$ (Thompson e Schumann 1987).

A força da evidência também pode ser resumida pela *razão de verossimilhança*:

$$\Lambda = \frac{P(E | C)}{P(E | I)}.$$

No exemplo, $\Lambda = 0,80/0,20 = 4$. A reação é quatro vezes mais provável entre culpados. Se $\Lambda = 1$, ela seria igualmente provável entre os dois tipos e não alteraria a suspeita.

Mesmo uma evidência forte precisa ser combinada com a probabilidade inicial. Bayes não apaga o ponto de partida. Voltaremos a essa observação ao discutir Red John.

4. O que é *screening*? O exemplo do seguro

Antes de retornar à série, vale estudar o *screening* em seu exemplo clássico: o mercado de seguros (Rothschild e Stiglitz 1976).

4.1. O problema da seguradora

Uma seguradora vende seguro de automóvel. Existem dois tipos de motorista:

- o cuidadoso, que sofre acidente com probabilidade 0,10;
- o arriscado, que sofre acidente com probabilidade 0,40.

Cada motorista conhece o próprio tipo, mas a seguradora não. Perguntar “você dirige com cuidado?” não resolve, pois todos gostariam de pagar o preço reservado ao motorista cuidadoso.

Se a empresa cobrar um único preço baseado no risco médio, o contrato ficará caro para o cuidadoso e barato para o arriscado. Alguns cuidadosos podem desistir, elevando o risco médio dos clientes que permanecem. Esse é um problema de *seleção adversa*, semelhante ao mercado de “limões” de Akerlof (1970).

4.2. A solução: oferecer duas opções

Em vez de pedir que o cliente declare seu tipo, a seguradora oferece dois contratos:

	Prêmio	Cobertura
Contrato A	alto	integral
Contrato B	baixo	franquia de R\$ 1.000

A franquia tem custo esperado diferente para cada tipo. Para o motorista cuidadoso, $0,10 \times 1000 = 100$. Para o motorista arriscado, $0,40 \times 1000 = 400$.

Chame de Δ a economia de prêmio obtida ao escolher o contrato B em vez do A. O cuidadoso prefere B quando a economia supera o custo esperado da franquia, isto é, quando $\Delta > 100$. O arriscado prefere A quando a economia oferecida por B não compensa seu custo esperado com a franquia, isto é, quando $\Delta < 400$. Portanto, se

$$100 < \Delta < 400, \quad (2)$$

o cuidadoso escolhe B e o arriscado escolhe A. A seguradora aprende algo sobre o risco do cliente observando o contrato escolhido.

Um exemplo numérico ajuda. Suponha que $\Delta = 250$. O cuidadoso economiza R\$ 250 no prêmio e assume custo esperado de R\$ 100: escolher B lhe dá vantagem esperada de R\$ 150. O arriscado também economiza R\$ 250, mas assume custo esperado de R\$ 400: para ele, B gera desvantagem esperada de R\$ 150. Assim, cada tipo escolhe um contrato diferente.

4.3. As três condições do mecanismo

O exemplo mostra três condições importantes.

Primeiro, o mesmo item precisa afetar os tipos de maneira diferente. A franquia custa mais, em valor esperado, para o motorista arriscado. Essa ordenação é chamada de *cruzamento único*.

Segundo, cada tipo deve preferir a opção desenhada para ele. Essa é a *compatibilidade de incentivos*. A desigualdade $100 < \Delta < 400$ garante que nenhum tipo queira imitar o outro.

Terceiro, os dois tipos precisam preferir algum contrato a ficar fora do mercado. Essa é a *restrição de participação*.

4.4. A informação tem custo

No mundo com informação completa, o motorista cuidadoso poderia receber cobertura integral por um preço baixo. No menu de *screening*, ele aceita uma franquía. Essa perda de cobertura não existe porque o cuidadoso gosta de risco, mas porque ajuda a impedir que o motorista arriscado imite sua escolha.

A separação, portanto, não é gratuita. A informação é produzida por meio de uma distorção. Períodos de experiência, garantias, versões gratuitas limitadas e prazos de carência usam lógica semelhante: opções distintas têm custos relativos diferentes para diferentes tipos.

4.5. Sinalização não é *screening*

Os conceitos são próximos, mas não iguais.

Quem move primeiro?

	Sinalização	<i>Screening</i>
Quem age primeiro?	O lado informado	O lado desinformado
Exemplo	Trabalhador escolhe estudar e apresenta diploma	Seguradora oferece um menu de contratos
Como surge a informação?	Pelo sinal escolhido pelo tipo	Pela opção escolhida dentro do mecanismo

Na sinalização de Spence (1973), o trabalhador conhece sua produtividade e escolhe apresentar um diploma. No *screening*, a seguradora não conhece o risco e desenha as opções. Em ambos os casos, a separação depende de custos diferentes entre tipos. A diferença está em quem desenha o movimento inicial.

5. O anúncio de Jane como *screening*

Voltemos à série. O anúncio de uma prova cria duas opções para o suspeito:

1. ficar parado e correr o risco de a prova aparecer;
2. reagir e tentar impedir a descoberta.

Jane é o lado desinformado: ele não conhece o tipo do suspeito. Mesmo assim, escolhe o anúncio, o momento e as pessoas que ouvirão a mensagem. Depois, observa as respostas. Essa sequência aproxima seu método do *screening*.

O mecanismo não é idêntico ao seguro. Não há contrato nem pagamento. O suspeito não escolhe participar do jogo. Além disso, a reação pode ocorrer sem ser observada. Ainda assim, a lógica econômica é a mesma: Jane modifica o problema de decisão para fazer com que tipos diferentes tenham incentivos diferentes. Em sentido mais amplo, ele desenha uma *estrutura informacional* (Bergemann e Morris 2019).

Há ainda um afastamento crucial em relação ao seguro: o menu não é anunciado como menu. Jane precisa que o suspeito interprete o anúncio como informação sobre uma prova, e não como uma armadilha construída para provocar uma escolha. A eficácia do mecanismo depende, portanto, de que ele não seja reconhecido como tal. A seção 9.1 mostra o que acontece quando Red John reconhece exatamente esse jogo.

A sequência do mecanismo pode ser resumida em cinco passos. Jane não observa quem é culpado. Ele anuncia que uma prova poderá revelar algo. Cada suspeito compara o custo de ficar parado com o

custo de reagir. Pessoas com perdas diferentes podem escolher ações diferentes. Por fim, Jane observa a resposta e atualiza sua suspeita.

O anúncio, isoladamente, pode ser apenas uma fala barata. Sua função não é convencer o público de que Jane sempre diz a verdade. Sua função é alterar os incentivos de quem escuta. A informação surge da ação provocada pela fala, não da fala em si.

6. O modelo básico: ficar parado ou reagir

Agora podemos traduzir a intuição em um modelo simples. Considere dois tipos de suspeito, culpado (C) e inocente (I), e duas ações:

$$a = 0 \text{ (ficar parado)}, \quad a = 1 \text{ (reagir)}.$$

Quatro parâmetros organizam a decisão.

Símbolo	Significado
q	crença do suspeito de que, se ficar parado, a investigação revelará a informação comprometedora
$L > 0$	perda sofrida caso essa informação seja revelada
$\sigma \in (0, 1]$	probabilidade de a tentativa de reagir ser detectada
$c > 0$	custo direto de reagir: esforço, tempo, dinheiro ou risco adicional

6.1. O custo de cada ação

Se o culpado fica parado, enfrenta perda esperada qL . Por exemplo, se acredita que a prova aparecerá com probabilidade 0,60 e a perda associada à descoberta é 100, sua perda esperada é $0,60 \times 100 = 60$.

Se reage, paga o custo c . Além disso, sua movimentação pode ser detectada com probabilidade σ , produzindo a perda L . Portanto, o custo esperado de reagir é $c + \sigma L$.

Para o inocente sem nada a ocultar, ficar parado não gera perda. Reagir custa c e não evita nada. As utilidades podem ser escritas como:

$$\begin{aligned} u(0 | C) &= -qL, & u(1 | C) &= -c - \sigma L, \\ u(0 | I) &= 0, & u(1 | I) &= -c. \end{aligned}$$

6.2. Derivando a condição de reação

O culpado reage quando a perda esperada de reagir é menor que a perda esperada de ficar parado:

$$c + \sigma L < qL.$$

Vamos reorganizar a desigualdade em três passos:

$$c + \sigma L < qL, \quad c < qL - \sigma L, \quad c < (q - \sigma)L.$$

Logo, a condição de reação é:

$$(q - \sigma)L > c. \tag{3}$$

Convém ler essa condição com atenção. A diferença $q - \sigma$ mede a parcela do risco de descoberta que o suspeito acredita poder evitar ao reagir. Ao multiplicar essa diferença por L , obtemos o benefício esperado da reação. O suspeito reage quando esse benefício supera o custo c .

A condição também mostra o efeito de cada parâmetro:

- q maior torna a reação mais provável: ficar parado parece mais perigoso;
- L maior torna a reação mais provável: há mais a perder;
- σ maior torna a reação menos provável: agir parece mais arriscado;
- c maior torna a reação menos provável: a tentativa fica mais cara.

6.3. Exemplo numérico

Suponha $q = 0,70$, $\sigma = 0,20$, $L = 100$ e $c = 30$. Então:

$$(q - \sigma)L = (0,70 - 0,20) 100 = 50.$$

Como $50 > 30$, o culpado reage. A mesma conclusão aparece ao comparar os custos: ficar parado custa $qL = 70$; reagir custa $c + \sigma L = 30 + 20 = 50$.

O inocente prefere ficar parado, pois $0 > -30$. Nesse caso, os dois tipos escolhem ações diferentes.

Proposição 1 (Separação entre culpado e inocente). *Se q , σ , L e c são conhecidos, existe separação estrita quando $(q - \sigma)L > c$: o culpado reage e o inocente fica parado. No modelo com apenas esses dois tipos, uma reação detectada implica $\mu(C | E) = 1$.*

Vale explicitar o que a armadilha identifica. No modelo mais simples, a armadilha identifica quem tem uma perda que pode ser evitada pela reação. Como o modelo supõe que apenas o culpado enfrenta essa perda, reagir e ser culpado coincidem. Essa conclusão deixará de valer quando permitirmos que um inocente também esconda algum segredo.

6.4. Agir não é o mesmo que ser observado

Uma pessoa pode reagir sem que Jane perceba. Se a probabilidade de detecção é $\sigma < 1$, a ausência de uma reação observada não prova que nada aconteceu.

Se $p = P(C)$ é a probabilidade inicial de culpa, a probabilidade posterior após nenhuma detecção é:

$$\mu(C | \neg E) = \frac{p(1 - \sigma)}{1 - p\sigma}. \quad (4)$$

Suponha $p = 0,25$ e $\sigma = 0,20$. Então:

$$\mu(C | \neg E) = \frac{0,25(0,80)}{1 - 0,25(0,20)} = \frac{0,20}{0,95} \approx 0,211.$$

A ausência de detecção reduz a suspeita de 25% para aproximadamente 21,1%, mas não a elimina.

6.5. O paradoxo da vigilância perfeita

Pode parecer que Jane deveria tornar a vigilância tão intensa quanto possível. O modelo mostra um limite. Se $\sigma = 1$ e todos sabem disso, qualquer reação será detectada com certeza. A condição de reação se transforma em $(q - 1)L > c$. Como $q \leq 1$ e $c > 0$, o lado esquerdo não pode superar o lado direito. Ninguém reage.

Corolário 1 (Vigilância perfeita e conhecida). *Se $\sigma = 1$ e esse valor é conhecido pelos suspeitos, a separação desaparece: reagir não é vantajoso para nenhum tipo.*

A armadilha precisa de vigilância suficiente para observar algumas reações, mas não tão evidente que torne a tentativa claramente inútil. Esse é um ponto importante: mais monitoramento pode aumentar a chance de detectar uma ação já escolhida, mas também pode reduzir a própria disposição de agir.

7. Uma reação não prova culpa

Na série, Jane às vezes encontra alguém que reagiu ao anúncio, mas escondia um segredo sem relação com o homicídio. O modelo básico não permite essa possibilidade porque divide todos em apenas dois grupos: culpado com algo a perder e inocente sem nada a ocultar.

Vamos ampliar os tipos: $\theta \in \{C, I_s, I_0\}$. Aqui:

- C é culpado do crime investigado;
- I_s é inocente do crime, mas esconde outro segredo;
- I_0 é inocente e não possui segredo relevante.

Cada tipo pode ter uma crença $q(\theta)$ e uma perda $L(\theta)$. A regra de decisão é a mesma:

$$[q(\theta) - \sigma]L(\theta) > c.$$

O culpado pode reagir para esconder o crime. O inocente com segredo pode reagir para esconder uma relação, uma fraude diferente, uma dívida ou qualquer outra informação que considere valiosa. O inocente sem segredo permanece parado.

Um exemplo com os três tipos torna a diferença visível. Imagine cem suspeitos: 25 culpados, 20 inocentes com outro segredo e 55 inocentes sem nada a ocultar. Suponha que os 25 culpados e os 20 inocentes com segredo reajam, enquanto os demais não reagem. Entre as 45 pessoas que reagiram, 25 são culpadas. Portanto,

$$P(C | E) = \frac{25}{45} = \frac{5}{9} \approx 0,556.$$

A reação ainda é informativa, pois elevou a suspeita de 25% para aproximadamente 55,6%. Mas está longe de provar culpa.

Formalmente, seja R o conjunto dos tipos que preferem reagir:

$$R = \{\theta \in \{C, I_s, I_0\} : [q(\theta) - \sigma]L(\theta) > c\}.$$

Se a mesma tecnologia de detecção vale para todos, então, após uma reação detectada:

$$\mu(C | E) = \frac{p(C) \mathbf{1}\{C \in R\}}{\sum_{\theta} p(\theta) \mathbf{1}\{\theta \in R\}}. \quad (5)$$

Proposição 2 (A partição produzida pela armadilha). *Se C e I_s preferem reagir, enquanto I_0 prefere ficar parado, a armadilha separa $\{C, I_s\}$ de $\{I_0\}$. Ela não separa $\{C\}$ de $\{I_s, I_0\}$. Assim, $\mu(C | E) = p(C)/[p(C) + p(I_s)] < 1$.*

Leitura econômica: a reação responde à pergunta “quem teme que a investigação revele algo?”, e não necessariamente à pergunta “quem cometeu este crime?”.

O inocente com segredo gera um *falso positivo*: o mecanismo o classifica como suspeito embora ele não tenha cometido o crime investigado. Também pode haver *falso negativo*: o culpado não reage, ou reage sem ser detectado.

Por isso, a ação observada deve orientar a investigação, não substituir a prova. Em linguagem institucional, o mecanismo serve para selecionar casos que merecem verificação adicional.

8. Credibilidade: o blefe de hoje enfraquece o de amanhã

Suponha que Jane anuncie uma prova decisiva em muitos casos e nunca apresente nada. Os suspeitos podem aprender com esse histórico. À medida que os anúncios falsos se acumulam, a crença q diminui.

Chame de β a fração observável de anúncios anteriores que se revelaram falsos. Podemos escrever $q = q(\beta)$, com $q'(\beta) < 0$. A margem líquida da reação é:

$$M(\beta) = [q(\beta) - \sigma]L - c. \quad (6)$$

O agente reage quando $M(\beta) > 0$. Como $q'(\beta) < 0$, mais blefes descobertos reduzem essa margem.

Um exemplo numérico mostra a força desse efeito. Considere $\sigma = 0,20$, $L = 100$ e $c = 30$. Se a credibilidade inicial é $q = 0,70$, então $M = (0,70 - 0,20)100 - 30 = 20$: vale a pena reagir. Se o histórico de blefes reduz q para $0,40$, então $M = (0,40 - 0,20)100 - 30 = -10$: o mesmo suspeito agora prefere ficar parado.

Proposição 3 (Autoconsumo da credibilidade). *Se $q'(\beta) < 0$, o conjunto de situações em que $M(\beta) > 0$ diminui à medida que o histórico de anúncios falsos cresce. Quando $q(\beta) \leq \sigma + c/L$, a condição de separação deixa de valer.*

Em jogos repetidos, as pessoas usam o histórico para avaliar quem faz o anúncio (Kreps e Wilson 1982; Milgrom e Roberts 1982). A lição é simples: credibilidade funciona como um estoque. Pode ser construída por promessas cumpridas e destruída por blefes revelados.

Isso cria uma escolha intertemporal. Um anúncio falso pode gerar informação hoje, mas reduzir a eficácia de anúncios futuros. Um investigador, regulador ou gestor precisa comparar o benefício imediato do blefe com a perda de reputação que ele produz.

9. Por que Red John é um adversário diferente?

Jane resolve muitos casos em poucos dias, mas leva anos para identificar Red John. O antagonista conhece o método, acompanha Jane e possui acesso a informações internas. Por isso, não enfrenta o mesmo problema de decisão dos suspeitos comuns.

9.1. Menor crença no anúncio e maior vigilância

Red John difere em duas dimensões. Primeiro, como conhece os blefes de Jane, atribui menor credibilidade ao anúncio: $q_{RJ} < q$. Segundo, por ser o principal alvo da investigação, enfrenta monitoramento mais intenso, de modo que a chance de uma movimentação ser detectada é maior: $\sigma_{RJ} > \sigma$.

Para os mesmos L e c , a margem de reação de um suspeito comum é $M = (q - \sigma)L - c$, enquanto a margem de Red John é $M_{RJ} = (q_{RJ} - \sigma_{RJ})L - c$. Subtraindo as duas expressões:

$$M - M_{RJ} = (q - \sigma)L - c - [(q_{RJ} - \sigma_{RJ})L - c] = [(q - q_{RJ}) + (\sigma_{RJ} - \sigma)]L.$$

Como $q - q_{RJ} > 0$ e $\sigma_{RJ} - \sigma > 0$, temos $M - M_{RJ} > 0$. Logo, $M_{RJ} < M$.

Corolário 2 (Adversário informado). *Se $q_{RJ} < q$ e $\sigma_{RJ} > \sigma$, então $M_{RJ} < M$. Se, além disso, $q_{RJ} < \sigma_{RJ} + c/L$, temos $M_{RJ} < 0$, e Red John prefere ficar parado.*

Leitura econômica: os dois efeitos caminham na mesma direção. Menor q_{RJ} reduz o risco que pode ser evitado; maior σ_{RJ} aumenta o risco de reagir. Conhecer o mecanismo torna mais fácil neutralizá-lo.

9.2. Quando exposição também é recompensa

Red John não busca apenas evitar punição. Ele também valoriza reconhecimento e notoriedade. O ponto de partida é a lógica de punição esperada de Becker (1968), acrescida aqui de uma recompensa pela atenção. A exposição pública, portanto, possui dois efeitos: aumenta a probabilidade de punição e aumenta a atenção recebida pelo agente.

Se π representa exposição, $\rho(\pi)$ é a probabilidade de punição, $R(\pi)$ é a notoriedade e $\gamma > 0$ mede quanto o agente valoriza reconhecimento, então:

$$U(\pi) = -\rho(\pi)L + \gamma R(\pi), \quad \rho'(\pi) > 0, \quad R'(\pi) > 0. \quad (7)$$

O efeito marginal da exposição é:

$$\frac{dU}{d\pi} = -L\rho'(\pi) + \gamma R'(\pi). \quad (8)$$

O primeiro termo é o freio: mais exposição aumenta a punição esperada. O segundo é o acelerador: mais exposição aumenta a recompensa simbólica.

Proposição 4 (Quando a exposição deixa de dissuadir). *Se $\gamma R'(\pi) > L\rho'(\pi)$, o ganho marginal de notoriedade supera o aumento marginal da perda esperada. Nesse caso, mais exposição eleva a utilidade do agente.*

A proposição não diz que publicidade sempre incentiva condutas ruins. Ela mostra que o efeito depende das preferências do agente. A mesma exposição pode desestimular quem teme reputação negativa e beneficiar quem transforma atenção em recompensa.

9.3. Bayes e a importância do ponto de partida

Suponha que uma evidência seja cem vezes mais provável se alguém for Red John do que se não for. Esse número parece enorme. Mesmo assim, o resultado final depende da probabilidade inicial. Se apenas uma pessoa em cem mil for Red John, podemos pensar em cem mil pessoas:

- uma é Red John e produz o sinal;
- entre as 99.999 restantes, uma taxa de falso positivo de 1% produz aproximadamente mil sinais.

Entre cerca de 1.001 sinais, apenas um pertence a Red John. A probabilidade posterior fica perto de 0,1%. Uma evidência forte pode ser insuficiente quando o tipo procurado é extremamente raro.

A lição prática é esta: nunca avalie uma pista apenas por sua aparência impressionante. Pergunte também: quão raro era o tipo antes de observar a pista? A evidência e o ponto de partida entram juntos na atualização bayesiana.

10. O mesmo mecanismo na economia real

O objetivo desta seção não é equiparar empresas, trabalhadores ou contribuintes a suspeitos. A analogia é estrutural: instituições frequentemente precisam aprender sobre características privadas e, para isso, desenham regras que induzem respostas diferentes.

Em cada aplicação, quatro perguntas são úteis:

1. Quem desenha o mecanismo?
2. Qual comportamento o mecanismo pretende induzir?
3. Que explicações legítimas podem gerar a mesma reação?
4. O anúncio continuará crível depois de ser usado repetidas vezes?

10.1. Fiscalização tributária, trabalhista ou ambiental

Uma autoridade pode anunciar que cruzará notas fiscais, declarações e movimentações bancárias de determinado setor. Algumas empresas se regularizam, corrigem erros ou organizam documentos. Outras cancelam operações ou transferem ativos.

Essas respostas contêm informação, mas não provam infração. Uma retificação pode ser tentativa de esconder evasão, mas também pode corrigir um erro genuíno.

No modelo, q é a crença de que a fiscalização realmente ocorrerá; L é a perda associada à irregularidade descoberta; c é o custo de corrigir, ocultar ou movimentar recursos; e σ é a chance de a movimentação ser percebida. Anúncios que nunca são executados reduzem q . Monitoramento claramente intenso aumenta σ . Os dois efeitos podem reduzir reações, embora por razões diferentes.

10.2. Auditoria e *compliance*

Imagine que uma empresa anuncie uma revisão dos pagamentos a fornecedores. Um gestor envolvido em fraude pode apagar uma planilha ou encerrar um contrato. Mas o setor financeiro também pode antecipar pagamentos para aproveitar descontos ou substituir um fornecedor por problemas de entrega.

Essa é a Proposição 2 fora da ficção. O anúncio pode separar quem teme alguma descoberta, mas não identifica automaticamente o motivo desse temor. A reação deve acionar uma segunda etapa: conferir contratos, autorizações, beneficiários finais e comunicações.

Documentação, direito de explicação e revisão independente reduzem o custo dos falsos positivos. Um bom mecanismo não apenas produz suspeitas; ele também cria meios de verificar se a classificação inicial estava correta.

10.3. Diplomas, selos e reputação em plataformas

Diplomas e certificações são exemplos de sinalização quando o lado informado escolhe apresentá-los. Limites graduais, garantias, pagamentos retidos e auditorias aleatórias são exemplos de *screening* quando a plataforma desenha as condições de participação.

Considere um selo de vendedor verificado que exige identidade confirmada, conta bancária compatível, entregas mínimas e poucas reclamações. Se vendedores confiáveis cumprem esses requisitos com mais facilidade, o selo informa. Se documentos e contas podem ser transferidos, vendedores oportunistas imitam os confiáveis e o valor do selo diminui.

Há ainda uma distinção importante: identidade autenticada não é qualidade comprovada. Saber quem controla a conta não garante que o produto seja autêntico ou que a entrega será realizada. Por isso, avaliações confirmadas, devoluções, disputas e histórico de transações precisam complementar a verificação inicial.

O desenho envolve uma troca. Exigências rigorosas dificultam a entrada de oportunistas, mas também podem excluir pequenos vendedores legítimos. Produzir informação custa alguma coisa, como vimos no exemplo da franquia do seguro.

10.4. Exposição pública e incentivos perversos

A publicidade de uma sanção pode elevar o custo reputacional de uma empresa, mas também ampliar a visibilidade de sua marca ou de sua campanha. Algo semelhante ocorre quando a cobertura de um ataque digital oferece reconhecimento ao autor.

O efeito líquido depende dos dois termos da equação $-L\rho'(\pi) + \gamma R'(\pi)$.

A Proposição 4 conclui diretamente sobre a utilidade do agente. Para passar dessa conclusão ao nível de atividade, é preciso supor que o agente dispõe de alguma margem de escolha: uma margem extensiva, entre participar ou não, ou uma margem intensiva, que permita escolher a intensidade de sua atuação.

Sob a hipótese adicional de que maior retorno líquido eleva a participação ou a intensidade, o aumento da punição esperada desestimula a atividade quando domina. Se o ganho de atenção domina, a divulgação pode produzir um incentivo perverso.

Isso não significa que a solução seja eliminar a transparência. Uma política mais cuidadosa pode divulgar o fundamento, o valor da sanção e as medidas corretivas sem reproduzir desnecessariamente a

mensagem ou a imagem que gera notoriedade.

Mapa das aplicações

Aplicação	Informação procurada	Principal risco
Fiscalização	Existência de irregularidade	Tratar correção legítima como confissão
Auditoria	Fraude ou conflito de interesse	Confundir prudência administrativa com culpa
Certificação	Qualidade ou confiabilidade	Imitação do selo e exclusão de bons agentes
Exposição pública	Resposta à punição reputacional	Transformar sanção em notoriedade

As aplicações convergem para a mesma lição. Boas instituições não se limitam a coletar dados. Elas desenham mecanismos críveis, observam respostas, verificam explicações alternativas e preservam a própria reputação ao longo do tempo.

11. Síntese do argumento

O método de Patrick Jane pode ser resumido em cinco movimentos:

1. existe uma característica privada que o investigador não observa;
2. o investigador altera o ambiente por meio de um anúncio;
3. cada tipo compara os custos de ficar parado e de reagir;
4. a escolha produz informação porque os incentivos diferem;
5. a nova evidência atualiza a suspeita, mas não substitui a investigação.

A condição $(q - \sigma)L > c$ resume o mecanismo básico. Ela diz que o suspeito reage quando o risco de descoberta que acredita poder evitar, multiplicado pelo tamanho da perda, supera o custo da reação.

O modelo também mostra quatro limites. Primeiro, a ausência de reação observada não prova inocência, pois uma tentativa pode não ser detectada. Segundo, uma reação não prova culpa, porque um inocente pode esconder outro segredo. Terceiro, o uso repetido de blefes reduz q e consome a credibilidade do instrumento. Quarto, a exposição pode perder seu efeito dissuasório quando o agente valoriza notoriedade.

A lição mais importante é que o mecanismo não lê mentes e não revela automaticamente a verdade. Ele transforma informação privada em comportamento observável. O comportamento altera probabilidades e ajuda a escolher onde investigar, mas ainda precisa ser confrontado com outras evidências.

12. Conclusão

The Mentalist pode ser lida como uma série sobre produção deliberada de informação. Patrick Jane não observa diretamente a culpa nem possui acesso à mente do suspeito. Ele constrói uma situação na qual agir e ficar parado têm consequências diferentes para pessoas diferentes.

Essa leitura aproxima a série de problemas centrais da economia da informação. Seguradoras oferecem menus de contratos; plataformas estabelecem selos, garantias e limites; autoridades anunciam fiscalizações; empresas criam auditorias. Em todos esses casos, a regra procura transformar uma característica escondida em uma escolha observável.

O principal cuidado é não pedir ao mecanismo mais do que ele pode entregar. Uma reação é uma evidência, não uma sentença. Seu valor depende da frequência de falsos positivos, da probabilidade inicial, da credibilidade do anúncio e dos incentivos específicos de cada agente. O bom desenho

institucional combina triagem com verificação posterior.

Patrick Jane, portanto, não lê mentes. Seu talento econômico é outro: ele cria escolhas capazes de produzir informação.

Referências

- Akerlof, George A. 1970. "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism." *Quarterly Journal of Economics* 84 (3): 488–500.
- Becker, Gary S. 1968. "Crime and Punishment: An Economic Approach." *Journal of Political Economy* 76 (2): 169–217.
- Becker, William E., and Michael Watts. 1996. "Chalk and Talk: A National Survey on Teaching Undergraduate Economics." *American Economic Review* 86 (2): 448–453.
- Bergemann, Dirk, and Stephen Morris. 2019. "Information Design: A Unified Perspective." *Journal of Economic Literature* 57 (1): 44–95.
- Crawford, Vincent P., and Joel Sobel. 1982. "Strategic Information Transmission." *Econometrica* 50 (6): 1431–1451.
- Farrell, Joseph, and Matthew Rabin. 1996. "Cheap Talk." *Journal of Economic Perspectives* 10 (3): 103–118.
- Ghent, Linda S., Alan Grant, and George Lesica. 2011. "The Economics of *Seinfeld*." *Journal of Economic Education* 42 (3): 317–318.
- Kreps, David M., and Robert Wilson. 1982. "Reputation and Imperfect Information." *Journal of Economic Theory* 27 (2): 253–279.
- Leet, Don, and Scott Houser. 2003. "Economics Goes to Hollywood: Using Classic Films and Documentaries to Create an Undergraduate Economics Course." *Journal of Economic Education* 34 (4): 326–332.
- Luccasen, R. Andrew, and M. Kathleen Thomas. 2010. "Simpsonomics: Teaching Economics Using Episodes of *The Simpsons*." *Journal of Economic Education* 41 (2): 136–149.
- Milgrom, Paul, and John Roberts. 1982. "Predation, Reputation, and Entry Deterrence." *Journal of Economic Theory* 27 (2): 280–312.
- Rothschild, Michael, and Joseph Stiglitz. 1976. "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information." *Quarterly Journal of Economics* 90 (4): 629–649.
- Spence, Michael. 1973. "Job Market Signaling." *Quarterly Journal of Economics* 87 (3): 355–374.
- Thompson, William C., and Edward L. Schumann. 1987. "Interpretation of Statistical Evidence in Criminal Trials: The Prosecutor's Fallacy and the Defense Attorney's Fallacy." *Law and Human Behavior* 11 (3): 167–187.
- Watts, Michael, and William E. Becker. 2008. "A Little More than Chalk and Talk: Results from a Third National Survey of Teaching Methods in Undergraduate Economics Courses." *Journal of Economic Education* 39 (3): 273–286.