How grower decision-making is modelled impacts the effect of plant disease control.

Rachel E. Murray-Watson *[1], Frédéric M. Hamelin[2], and Nik J. Cunniffe[1]

[1]Department of Plant Sciences, University of Cambridge, Cambridge, United Kingdom

[2]IGEPP, INRAE, Institut Agro, Univ Rennes, Rennes, France

Short title: Human behaviour and plant disease management.

***For correspondence:** rm844@cam.ac.uk

Abstract

While the spread of plant disease depends strongly on biological factors driving transmission, it also has a human dimension. Disease control depends on decisions made by individual growers, who are in turn influenced by a broad range of factors. Despite this, human behaviour has rarely been included in plant epidemic models. Considering Cassava Brown Streak Disease, we model how the perceived increase in profit due to disease management influences participation in clean seed systems (CSS). Our models are rooted in game theory, with growers making strategic decisions based on the expected profitability of different control strategies. We find that both the information used by growers to assess profitability and the perception of economic and epidemiological parameters influence long-term participation in the CSS. Over-estimation of infection risk leads to lower participation in the CSS, as growers perceive that paying for the CSS will be futile. Additionally, even though good disease management can be achieved through the implementation of CSS, and a scenario where all controllers use the CSS is achievable when growers base their decision on the average of their entire strategy, CBSD is rarely eliminated from the system. These results are robust to stochastic and spatial effects. Our work highlights the importance of including human behaviour in plant disease models, but also the significance of how that behaviour is included.

1 Introduction

Human behaviour is intuitively important in the control of plant infectious disease, as individuals often face choices when adopting infection-limiting behaviours. This behaviour has been widely considered in the context of human (e.g. [1], [2], [3], [4], [5]; reviewed in [6] and [7]) and animal (e.g. [8], [9],[10], [11]) diseases, though there have been fewer studies for plant diseases (with the exception of [12], [13], [14], [15], [16] and [17]). Where studies do consider individual's decisions there is much variation in how decision-making is modelled, although it clearly will affect the outcomes of their decisions.

We assume that decisions regarding disease control are intrinsically linked to disease prevalence. Clearly, they must also depend upon considerations such as the cost of control, the consequences of infection and the perceived risk of becoming infected ([18], [19], [20]). Additionally, and crucially, decisions made by one grower will influence the decisions of another. Voluntary disease control programmes can therefore be viewed as a collective action problem (also termed social dilemmas; [21]) as they generate externalities that affect the outcomes of those not involved in the programme. By reducing or eliminating their own chance of infection, growers who control also lower the probability that their neighbours will be infected, thereby generating a positive externality that disincentives their neighbours' engagement in disease prevention [22].

Such "free-riding" behaviour is indirectly encouraged when there is higher overall participation in control schemes, as growers are less likely to become infected whenever the proportion controlling is high. Yet if too many growers engage in free-riding behaviour, there will be a resurgence of infections, which may make disease control more likely. These feedback loops are believed to be a major reason why some public health schemes for human disease, such as subsidised vaccination campaigns, have failed to lead to disease eradication [22]. Consequently, the success of voluntary control policies may be self-limiting. Indeed, perceived risk of other growers benefiting from an individual's control policies are a prominent reason why growers do not participate in control schemes ([18], [23]).

The strategic nature of grower behaviour can be modelled using game-theoretic analysis. Game theory is used to study decision making in settings where an individual (or "player") can choose from a variety of strategies, with the aim of choosing one that will maximise their desired outcome ("payoff") [24]. No individual makes a decision in isolation, and the best strategy to play will depend on the strategy choices of others. In our work, we assume growers have access to different combinations of three pieces of information: their own yield from the previous year, an estimate of the expected profits of each strategy (control or non-control) and the overall expected profit of the population. They then compare a subset of these different quantities to assess profitability and thus whether they should change strategy. Though the assumption that growers have access to "perfect

information" is limiting and unlikely to bear true in a real-world context, understanding the results in such a case is still a helpful tool in studying the effects of these assessments on participation in control. The case of "perfect information" also acts as a useful baseline from which other scenarios where growers have less information may be assessed.

Game theoretic approaches have been used in previous disease models, often based on evolutionary game theory (e.g.[25] and [10]; for non-epidemic uses see: [26], [27], [28], [29], [30], [31]). In this formalism, in a two-strategy model, the performance of the player's own strategy is compared against that of the average performance of the population. If the payoff is less than that of the population, the focal player changes strategy; if it is higher, they remain with their current strategy. For plant disease, this would involve growers comparing the average profit of all of those using their own control strategy with the average profit over all growers in the population (we call this the "strategy vs. population" model). Of course this assumes, perhaps unrealistically, that growers will base their decision making upon the profit of all those using the same control strategy.

An alternative means of assessing profitability, is used in [12], [13] and [17] (though the inclusion of grower characteristics in [13] means it deviates from pure game-theoretic assumptions of rationality as growers may adopt a strategy that will earn them lower profits). In these two-strategy models, growers compare their own profit with the average profit of those playing the alternative strategy (the strategy that the grower does not currently use; for a controller, this would be non-control and vice versa). We henceforth call this the "grower vs. alternative" model. Ostensibly this seems a more likely comparison a grower will make, as it requires less information and focuses more closely on a grower's own performance, rather than all who act similarly to that grower. Other possible models combine elements of both approaches, since growers compare their strategy with the expected profit of the alternative strategy ("strategy vs. alternative" model) or compare their own profit with the average of the population ("grower vs. population" model). To our knowledge, the "grower vs. population" and "grower vs. alternative" models have never been compared.

The two model classes ("strategy vs." and "grower vs.") can be broadly classified based on the

economic concepts of "rational" and "adaptive" expectations, respectively ([32]). If an individual forms their expectations rationally, they assess the information available at any instant and then extrapolate/determine what that means for the future. This has a clear relationship with our "strategy vs." models, which use the current probabilities of infection as a proxy for what the future probabilities will be. Conversely, adaptive expectations are based solely on historical outcomes. Rather than simply being "adaptive", our "grower vs." models can instead be viewed as "strategic-adaptive" ([33]), accounting for prior experience whilst still accounting for future events.

Hypothetical scenarios in which each of these model types could potentially apply can be constructed. If a grower is part of an area-wide control scheme (such as the citrus health management areas (CHMAs) for citrus diseases in the United States of America; [34], or to manage Queensland fruitfly populations in Australia, [35]) that only includes a subset of the population, they may wish to compare the profits of the scheme with the expected profit of the entire population (the "strategy vs. population" comparison). The "strategy vs. alternative" comparison may be used if a grower is part of an area-wide control scheme but wishes to join another scheme with an different means of control. If a particular strategy may confer a market advantage to the grower (for example, if pesticide use is widespread, they may wish to stop spraying to then market themselves as "organic"), the grower may compare their outcome with the expected profit of the population ("grower vs. population"). Finally, if a grower is using a particular control strategy and then approached by an extension worker who proposes an different strategy, they may employ the "grower vs. alternative" comparison.

To examine the effects of these different models of growers' decision making in a concrete setting, we use a simple model describing the spread of Cassava Brown Streak Disease (CBSD) based on the model presented in [13]. Cassava (*Manihot esculenta*) is a staple food in Sub-Saharan Africa, where it is predominantly grown as a subsistence crop. Cassava production is threatened by CBSD, a viral disease caused by either cassava brown streak virus (CBSV, [36]) or Ugandan CBSV (UCBSV; [37]). Infection results in necrosis of the stems and tubers, leading to yield losses of up to 70%

[38]. Coupled with its increasing spread from east to west Africa, this positions CBSD as a major challenge to cassava cultivation across the region [39].

Both viruses are transmitted horizontally by a whitefly vector (*Bemisia tabaci*, [40]) and vertically through the replanting of infected stem cuttings. Informal trade of cassava cuttings is widespread amongst growers, though the subtlety of symptoms often means that the practice often contributes to vegetative propagation of infected material [41] and spread of CBSD between growers [42].

These potential for vertical transmission means that clean seed systems (CSSs) are a plausible CBSD management strategy [43]. CSSs disseminate virus-free planting material, thereby avoiding the vegetative propagation of infected material. The success of CSSs will depend on grower behaviour, as participation will be governed by the balance of costs and benefits of the scheme. This, and the combination of trade- and whitefly-mediated pathogen transmission, means that a grower's profit will depend on the action of others.

The overall objective of this work is to investigate how including grower behaviour in plant epidemic models affects the outcome of control schemes, and how the factors influencing behaviour can best be manipulated to encourage control uptake. We use our behavioural model alongside the CBSD case study to address the following: (1) How does profit comparison's formulation (i.e. "strategy vs." or "grower vs.") affect the model and its results, and can a scenario where all growers use the CSS be attained? (2) How does grower participation in the CSS depend on epidemiologically and economically important parameters, and how can high levels of control be encouraged? (3) How does systemic uncertainty in epidemiological parameters affect participation in the CSS? Though the "strategy vs." models are mathematically tractable and therefore allow for the derivation of analytical expressions, our view is that it is less likely that growers will make their assessment of profitability based on every other grower using the same strategy as they are, rather than just their own outcomes. Thus, after our initial comparisons (Question 1 above), we focus solely upon the "grower vs. alternative" model to investigate the effect of parameters and systematic uncertainty on CSS participation (Questions 2 and 3 above).

2 Materials and Methods

2.1 Epidemiological model

Our model (Figure 1(A)) is a simplified representation of a CSS. Fields are classified by infection status (susceptible, S, or infected, I), as well as whether their growers currently control, i.e. used certified clean seed the last time the field was planted (subscripts: controller, C, or non-controller, N). We assume each grower cultivates only a single field, and so use the terms "grower" and "field" synonymously. Infection status and control status are binary classifications; we do not model within-field spread of disease, and assume controllers plant only certified virus-free material.

Susceptible fields are vulnerable to horizontal infection (at rate β) due to viruliferous vectors moving from infected fields (fields of class I_C and I_N) (Table 1). Vertical infection also occurs via infected propagation material. We assume that the propagation material used by controllers is never infected. The probability of a non-controller becoming vertically infected depends jointly on the probability of acquiring infected planting material from an individual infected field (p) and the prevalence of disease at the time of planting $\left(\frac{I_C+I_N}{N}\right)$. The season length ($1/\gamma$) is independent of the control strategy adopted in, and infection status of, any given field. Since we use a continuous time model, fields are asynchronously harvested and replanted [44], a plausible assumption for cassava cultivation in much of Sub-Saharan Africa [45].

To integrate behavioural dynamics, we represent the decisions made by growers at the time of planting by switching terms. These reflect the probability of growers switching strategy for their following crop, based on their current strategy and the current state of the system, as well as, potentially, whether their previous crop was infected. In general, we distinguish four switching

Table 1: Summary of parameter values.

Parameter	Meaning	Value	Reference
$1/\gamma$	Length of the growing season	300 days	[13]; [46]
β	Rate of secondary infection	8.93×10^{-6} day^{-1} field^{-1}	Calibrated to [13]
p	Probability of planting infected cuttings	0.8	Assumed (see main text)
η	Responsiveness of growers	10	Assumed (see main text)
Y	Maximum yield	1	All values scaled relative to yield
L	Loss due to infection	0.6	[38]; [47]
ϕ	Cost of control	0.25	[13]
N	Total number of fields/growers	750	Illustrative
$S_C(0)$	Initial proportion of susceptible controllers	$0.1N$	Illustrative
$I_C(0)$	Initial proportion of infected controllers	0	Illustrative
$S_N(0)$	Initial proportion of susceptible non-controllers	$0.89N$	Illustrative
$I_N(0)$	Initial proportion of infected non-controllers	$0.01N$	Illustrative

terms (Figure 1(A)).

$$z_{SC} = \mathbb{P}\left(\text{switch from } C \text{ to } N | \text{previous crop remained uninfected } (S)\right), \tag{1}$$

$$z_{IC} = \mathbb{P}\left(\text{switch from } C \text{ to } N | \text{previous crop was infected } (I)\right), \tag{2}$$

$$z_{SN} = \mathbb{P}\left(\text{switch from } N \text{ to } C | \text{previous crop remained uninfected } (S)\right), \tag{3}$$

$$z_{IN} = \mathbb{P}\left(\text{switch from } N \text{ to } C | \text{previous crop was infected } (I)\right). \tag{4}$$

As described below, the probabilities encoded in the switching terms depend on growers' assessments of likely profits, and thus depend on the comparisons we assume are made by growers.

Our general model is

$$\frac{dS_C}{dt} = \gamma\theta_C - \beta S_C\left(I_C + I_N\right) - \gamma S_C, \tag{5}$$

$$\frac{dI_C}{dt} = \beta S_C\left(I_C + I_N\right) - \gamma I_C, \tag{6}$$

$$\frac{dS_N}{dt} = \gamma\theta_N\left(1 - \frac{p(I_C + I_N)}{N}\right) - \beta S_N\left(I_C + I_N\right) - \gamma S_N, \tag{7}$$

$$\frac{dI_N}{dt} = \gamma\theta_N\left(\frac{p(I_C + I_N)}{N}\right) + \beta S_N\left(I_C + I_N\right) - \gamma I_N. \tag{8}$$

where:

$$\theta_C = S_C(1 - z_{SC}) + I_C(1 - z_{IC}) + S_N z_{SN} + I_N z_{IN} \tag{9}$$

$$\theta_N = S_C z_{SC} + I_C z_{IC} + S_N(1 - z_{SN}) + I_N(1 - z_{IN}) \tag{10}$$

2.2 Parameterisation

Baseline parameter values for CBSD were taken from previous studies where possible (Table 1). The rate of secondary infection, β, was set by matching the results of our model against [13]. In particular, we used the baseline behaviour of that model to motivate the choice that, in the absence of trade-mediated transmission, with 1% of fields initially infected and with no option of controlling for disease, 50% of fields would be infected within 10 seasons. To only account for the horizontal transmission, the probability of vertical transmission, (p) was set to zero when matching these results.

The value of 60% of fields' potential profit lost due to infection was assumed, to account for both the loss of starch content – which causes up to 40% loss of yield [47] – and the reduced market value of infected materials [38].

Vertical transmission requires cuttings taken from infected plants to be replanted and cause infection in the next season. The probability of vertical transmission is therefore reduced by selection,

in which growers avoid replanting visibly infected stems [48] and by reversion, in which low viral titres mean healthy cuttings can be taken even from infected plants [49], a process that has been acknowledged for a range of virus diseases [50]. It is difficult to quantify these effects into a single parameter value; we therefore take $p = 0.8$ as a pragmatic but arbitrary default in which vertical transmission at the field scale often, but not always, occurs.

Growers' responses to differences in profit are notoriously difficult to quantify [15], and here we were forced to assume a value for our parameter describing the responsiveness of growers (η). An intuition for our default selection of $\eta = 10$ per unit of profit follows by noting that, using the default costs and losses in the"strategy vs. population" model, a controller will have an approximately 80% chance of switching to become a non-controller if there is sufficient infection in the rest of the system such that it is almost certain they will become infected over the next season (i.e. if the probability of infection of a controller, $q_C, \approx 1$; Equation 17).

2.3 Strategies, outcomes and profits

For simplicity, we assume only a single control strategy is available, the CSS. Thus, there are only two pure strategies available to growers at the beginning of each season: to control for disease, or to not control (Figure 1(B)). Within each strategy, there are two outcomes realised at the time of harvesting: the grower's field may remain uninfected or have become infected.

Each strategy $\times$ outcome combination results in a different payoff (Figure 1(B)). An uninfected field generates yield Y, with loss of income, L, in infected fields. Control incurs a cost ϕ. The profit

for each outcome is (Table 3):

$$P_{SN} = \text{Profit for non-controller who remains susceptible} = Y, \tag{11}$$

$$P_{IN} = \text{Profit for non-controller who becomes infected} = Y - L, \tag{12}$$

$$P_{SC} = \text{Profit for controller who remains susceptible} = Y - \phi, \tag{13}$$

$$P_{IC} = \text{Profit for controller who becomes infected} = Y - \phi - L, \tag{14}$$

Rational growers will never control if $L < \phi$, since then the cost of clean seed exceeds the loss due to infection. We therefore only consider $L > \phi$, from which it follows that:

$$P_{SN} > P_{SC} > P_{IN} > P_{IC}. \tag{15}$$

2.4 Risks of infection

We initially focus on the case in which growers can estimate instantaneous risks of infection precisely. In particular, we assume that growers use their current instantaneous probability of infection as an estimate of their instantaneous probability of horizontal infection ($p(\text{Horiz})$) over the next season:

$$\begin{aligned} p(\text{Horiz}) &= \frac{\text{Instantaneous infection rate}}{\text{Instantaneous infection rate + Harvesting rate}}, \\ &= \frac{\beta(I_C + I_N)}{\beta(I_C + I_N) + \gamma}. \end{aligned} \tag{16}$$

For controllers, clean seed rules out vertical transmission. Equation 16 therefore sets a grower's estimate of the probability of infection next season if they were to choose to control (q_C):

$$\begin{aligned} q_C &= \text{Grower's estimate of the probability of infection next season if control is adopted}, \\ &= \frac{\beta(I_C + I_N)}{\beta(I_C + I_N) + \gamma}. \end{aligned} \tag{17}$$

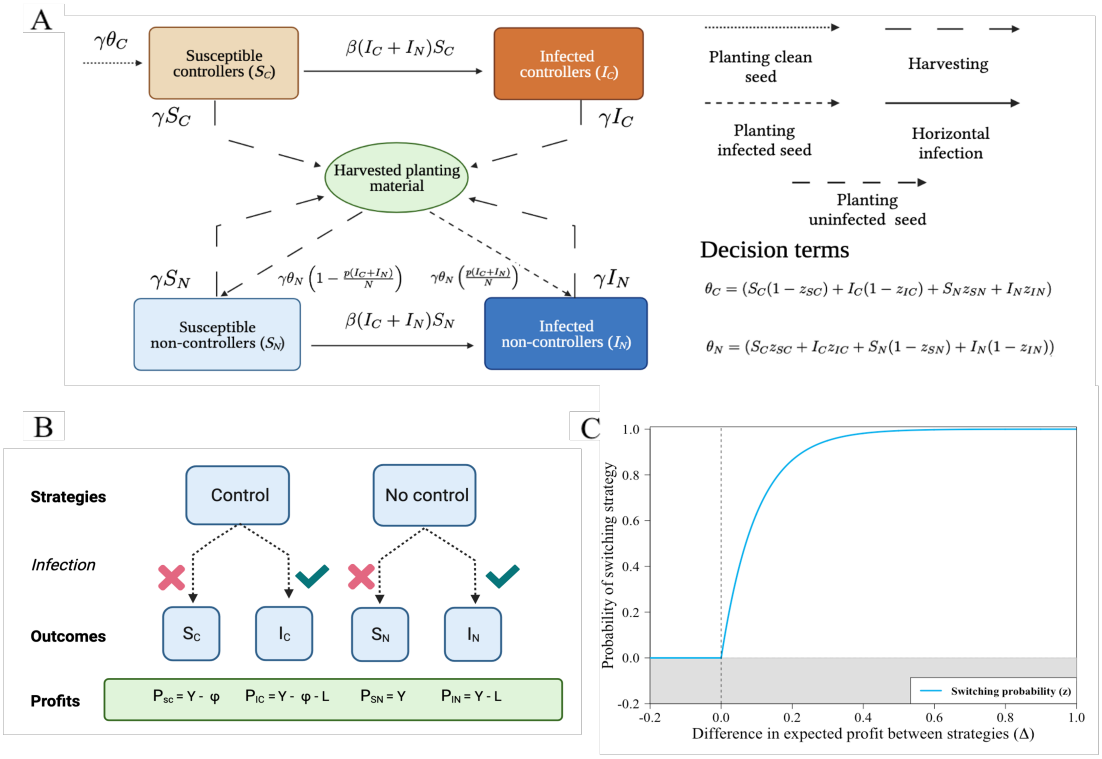

Figure 1: Epidemiological and behavioural models. (A) The epidemiological model distinguishes growers who control (i.e. plant clean seed) from those who do not. Only non-controllers can become infected via vertical transmission, although all growers are subject to horizontal transmission due to movement of vectors from infected fields (Equations 5-8). When replanting growers potentially switch strategies. (B) Strategies, outcomes and profits; the latter account for the loss of yield if infected and the cost of participating in the control scheme. (C) Decisions to switch are based on a switching function, parameterising the probability of switching strategy z as a function of Δ as difference in predicted profits. Figures (A) and (B) were created with BioRender.com.

For non-controllers, vertical infection must also be considered. We again assume growers can estimate the relevant instantaneous probability of vertical transmission ($p_N(\text{Vert})$) with perfect accuracy:

$$\begin{aligned} p_N(\text{Vert}) &= \frac{\text{Probability of vertical transmission}}{\text{Total number of fields}}, \\ &= \frac{p(I_C + I_N)}{N}. \end{aligned} \tag{18}$$

A non-controller's instantaneous probability of horizontal infection ($p(\text{Horiz})$) must account for the fact that they may first be infected via vertical transmission, leading to:

$$\begin{aligned} p_N(\text{Horiz}) &= \begin{matrix}\text{Instantaneous probability} \\ \text{of horizontal transmission}\end{matrix} \times \left(1 - \begin{matrix}\text{Instantaneous probability} \\ \text{of vertical transmission}\end{matrix}\right), \\ &= \left(\frac{\beta(I_C + I_N)}{\beta(I_C + I_N) + \gamma}\right)\left(1 - \frac{p(I_C + I_N)}{N}\right). \end{aligned} \tag{19}$$

Combining these gives us the instantaneous probability of infection for a non-controller (q_N):

$$\begin{aligned} q_N &= \text{Grower's estimate of the probability of infection next season if control is not adopted}, \\ &= p_N(\text{Vert}) + p_N(\text{Horiz}), \\ &= \frac{p(I_C + I_N)}{N} + \left(\frac{\beta(I_C + I_N)}{\beta(I_C + I_N) + \gamma}\right)\left(1 - \frac{p(I_C + I_N)}{N}\right). \end{aligned} \tag{20}$$

How these probabilities of infection change with increasing disease pressure ($I_N + I_C$) is shown in Appendix 1 Figure 1 (A).

2.5 Expected profits

Our behavioural models assume that – at the time of planting – individual growers estimate the expected profit for the next season of adopting each strategy. These depend on the payoffs for each strategy × outcome combination (Equations 11-14) and the estimated probabilities of infection under

each strategy (Equations 17 and 20). In particular, the estimated expected profit for a controller, P_C, would be

$$\begin{aligned} P_C &= \text{Grower's estimate of the expected profit next season if control is adopted,} \\ &= q_C P_{IC} + (1 - q_C) P_{SC}, \\ &= Y - \phi - L\left(\frac{\beta(I_C + I_N)}{\beta(I_C + I_N) + \gamma}\right). \end{aligned} \tag{21}$$

The corresponding estimate without control, P_N, is

$$\begin{aligned} P_N &= \text{Grower's estimate of the expected profit next season if control is not adopted,} \\ &= q_N P_{IN} + (1 - q_N) P_{SN}, \\ &= Y - L\left(\frac{p(I_C + I_N)}{N} + \left(\frac{\beta(I_C + I_N)}{\beta(I_C + I_N) + \gamma}\right)\left(1 - \frac{p(I_C + I_N)}{N}\right)\right). \end{aligned} \tag{22}$$

The expected profit averaged over all growers, P, is

$$\begin{aligned} P &= \text{Grower's estimate of the expected profit next season averaged over the population,} \\ &= P_C\left(\frac{S_C + I_C}{N}\right) + P_N\left(\frac{S_N + I_N}{N}\right). \end{aligned} \tag{23}$$

Calculating P therefore requires growers to know the instantaneous proportion of the population controlling. This calculation of P is instantaneous, i.e. based on the state of the system at the time of planting, and provides an estimate for the next season.

2.6 Behavioural models

We systematically examine different ways to construct the switching terms in Equations 5-8. Differences between behavioural models turn on how Δ, an expected difference in profit (and thus profitability), is estimated. We consider four possible comparisons:

- *Strategy vs. population.* Growers compare the expected profit for their current strategy with the expected profit across the entire population of growers ($P - P_j$, with $j \in \{C, N\}$).

- *Strategy vs. alternative strategy.* Growers compare the expected profit for their current strategy with the expected profit of the strategy which they did not use ("alternative strategy"; $P_i - P_j$, with $i, j \in \{C, N\}$ and $i \neq j$).

- *Grower vs. population.* Growers compare their own outcome during the previous season with the expected profit across the entire population of growers ($P - P_G$, where P_G is the grower's profit from the previous season).

- *Grower vs. alternative strategy.* Growers compare their own outcome during the previous season with the expected profit of the alternative strategy ($P_i - P_G$, where $i, j \in \{C, N\}$).

All set z, the probability of an individual grower switching strategy, as a function of Δ, via

$$z = \max\left(0, 1 - e^{-\eta\Delta}\right) = \begin{cases} 1 - e^{-\eta\Delta} & \text{if } \Delta \geq 0 \\ 0 & \text{otherwise.} \end{cases} \tag{24}$$

Since growers aim to maximise their profit, they only consider switching strategy when the comparison appears to offer a higher payoff (i.e. when $\Delta > 0$; Figure 1(C)). The parameter η sets the responsiveness of growers to a unit difference in profit.

A summary of the information required for each means of comparison is provided in Table 2.

The "strategy vs." models are inspired by evolutionary game theory ([25] and [10]), with Δ depending on the expected profit for the strategy currently adopted. The "grower vs." models are closer to previous usage in plant disease epidemiology (e.g. [13], [12] and [17]), with the outcome obtained by the grower for their last crop used as the point of comparison. Table 3 provides a summary of the symbols used in the behavioural model.

Table 2: Summary of the information required by growers to carry out each means of comparison. "Controller prevalence" refers to the proportion of growers that adopted control the last time their field was planted (i.e. $\frac{S_C+I_C}{N}$) and "disease prevalence" is the proportion of growers that are currently infected (i.e. $\frac{I_N+I_C}{N}$). The "✓" refers to when information is required, and "X" when it is not necessary for decision-making.

	Disease prevalence	Controller prevalence	Own outcome
Strategy vs population	✓	✓	X
Strategy vs alternative	✓	X	X
Grower vs population	✓	✓	✓
Grower vs alternative	✓	X	✓

Table 3: Summary of symbols used in defining the behavioural models. The outcomes $i \in \{S, I\}$ represent – at the end of the season – fields that are (S)usceptible and (I)nfected, respectively. The strategies $j \in \{C, N\}$ correspond to (C)ontrollers (i.e. use clean seed) and (N)on-controllers (i.e. do not use clean seed), respectively.

Symbol	Meaning	Defined in
z_{ij}	Switching probability for grower in infection class $i \in \{S, I\}$ who used strategy $j \in \{C, N\}$	Equations 1-4
P_{ij}	Profit for grower with outcome $i \in \{S, I\}$ who used strategy $j \in \{C, N\}$	Equations 11-14
q_j	Estimated probability at time of planting of infection if go onto use strategy $j \in \{C, N\}$	Equations 17 & 20
p_N(Vert)	Estimated probability of vertical infection for non-controllers	Equation 18
p_N(Horiz)	Estimated probability of horizontal infection for non-controllers	Equation 19
P_j	Estimated profit at time of planting if go onto use strategy $j \in \{C, N\}$	Equations 21 and 22
P	Estimated profit at time of planting averaged over entire population	Equation 23
Δ	Estimated difference in expected profit	Equation 24
q_β	Perceived value of the rate of horizontal transmission (β)	Equation 37
q_I	Perceived number of infected fields	Equation 38

2.6.1 Strategy vs. population

In the "strategy vs." models, the switching probabilities are decoupled from the outcome during the previous season, i.e. $z_{SC} = z_{IC} = z_C$ and $z_{SN} = z_{IN} = z_N$. The probability of a controller switching to no longer use clean seed is

$$z_{SC} = z_{IC} = z_C = \max\left(0, 1 - e^{-\eta(P-P_C)}\right). \tag{25}$$

This probability is only non-zero when the expected profit for a controller is assessed to be smaller than the expected profit over the entire population (i.e. when $P_C < P$). Conversely, the probability of a non-controller switching is

$$z_{SN} = z_{IN} = z_N = \max\left(0, 1 - e^{-\eta(P - P_N)}\right), \tag{26}$$

which takes a non-zero value only when $P_N < P$. Since P is a convex combination of P_C and P_N (the two bracketed terms in Equation 23 are complementary probabilities), only one of P_C or P_N can be greater than P at any time, meaning z_C and z_N cannot simultaneously be non-zero (although if $P_C = P_N$, i.e. the expected profits are equal, then $P = P_C = P_N$ and so $z_C = z_N = 0$).

How the expected profits and switching terms change with increasing disease pressure $(I_N + I_C)$ is shown in Appendix 1 Figure 1(B)-(C) and (E)-(F).

2.6.2 Strategy vs. alternative strategy

This model is similar to the "strategy vs. population" model, but now instead of comparing to the expected profit of the population, P, growers compare to the expected profit of the alternative strategy (i.e. P_C or P_N, as appropriate). The probability of a controller switching is

$$z_{SC} = z_{IC} = z_C = \max\left(0, 1 - e^{-\eta(P_N - P_C)}\right), \tag{27}$$

which takes a non-zero value only when $P_N > P_C$, i.e. when non-controllers are expected to be more profitable next season. Conversely, the probability of a non-controller switching is

$$z_{SN} = z_{IN} = z_N = \max\left(0, 1 - e^{-\eta(P_C - P_N)}\right), \tag{28}$$

which is non-zero only when $P_C > P_N$. Here it is clear that only one of z_C and z_N can be non-zero at any time, since either $P_C > P_N$ or $P_C < P_N$ (or $P_C = P_N$, in which case no grower would switch

since the expected profits of the two strategies are equal, and $z_C = z_N = 0$).

2.6.3 Grower vs. population

In the "grower vs." models, decisions are made on the basis of grower's own outcome for the last crop, and so all four switching probabilities can differ. The ordering of profits in Equation 15 means that an uninfected non-controller (i.e. class S_N) should never switch strategy, since these "successful free-riders" [51] obtained the highest possible profit for the last crop. Similarly, having obtained the "suckers' pay off", a grower who controlled for disease but was nevertheless infected (i.e. class I_C) will always have a non-zero probability of switching, since they must have obtained a lower than average profit. Whether or not S_C and I_N growers switch strategy depends on their profit relative to the expectation over the whole population, P. The switching terms are therefore

$$z_{SN} = 0, \tag{29}$$

$$z_{IN} = \max\left(0, 1 - e^{-\eta(P - P_{IN})}\right), \tag{30}$$

$$z_{SC} = \max\left(0, 1 - e^{-\eta(P - P_{SC})}\right), \tag{31}$$

$$z_{IC} = 1 - e^{-\eta(P - P_{IC})}. \tag{32}$$

2.6.4 Grower vs. alternative strategy

This model is similar to the "grower vs. population" model, but now growers compare to the expected profit of the alternative strategy (i.e. P_C or P_N) rather than that of the population, P,

with

$$z_{SN} = 0, \tag{33}$$

$$z_{IN} = \max\left(0, 1 - e^{-\eta(P_C - P_{IN})}\right), \tag{34}$$

$$z_{SC} = \max\left(0, 1 - e^{-\eta(P_N - P_{SC})}\right), \tag{35}$$

$$z_{IC} = 1 - e^{-\eta(P_N - P_{IC})}. \tag{36}$$

2.7 Systematic uncertainty

The models as presented thus far presume that all growers can perfectly perceive the risk of infection by knowing the true values of β and the total number of infected individuals (I). However, these quantities must actually be estimated by growers. To account for this, we introduce two new parameters: the perceived rate of horizontal transmission (q_β) and the perceived number of infected fields (q_I). These are given by:

$$q_\beta = \nu_\beta \beta, \tag{37}$$

$$q_I = \nu_I (I_N + I_C). \tag{38}$$

Therefore ν_β and ν_I scale the values of β and $(I_C + I_N)$ respectively, amounting to either an over- or underestimate of their values. This modifies their contributions to the estimated profits in Equation 21 and Equation 22, which become:

$$\tilde{P}_C = Y - \phi - L\left(\frac{q_\beta q_I}{q_\beta q_I + \gamma}\right). \tag{39}$$

and

$$\tilde{P}_N = Y - L\left(\frac{pq_I}{N} + \left(\frac{q_\beta q_I}{q_\beta q_I + \gamma}\right)\left(1 - \frac{pq_I}{N}\right)\right). \tag{40}$$

Importantly, these changes to how growers perceive the risk of infection do not change the epidemiology of the system, only the behaviour of the growers. If the value of $\nu_I(I_N + I_C) > N$, we assume that growers estimate that all fields in the system are infected and thus set $\nu_I(I_N + I_C) = N$.

2.8 Spatial-stochastic model.

We use an individual-based version of our model to test the robustness of our conclusions to spatially explicit and/or stochastic effects (see Appendix 2 for full details). Essentially, we modelled disease spread on a 1,000 Ha landscape containing $N = 750$ fields. Disease spreads both via movement of a whitefly vector (with the probability of a field becoming infected depending on its proximity to other infected fields) and via replanting of infected material. Decisions regarding control (which take the same functional form as the deterministic models, but now account for the dependence of the force of infection on each grower's spatial location) are made at the time of planting, which occurs asynchronously across fields. We also investigated the effect of growers' misperception in this model, focusing on the dispersal scale of the whitefly vector (α) and the rate of horizontal transmission β_S (Appendix 2 Table 1).

3 Results

3.1 Effect of profit comparison's formulation.

3.1.1 Characterising model equilibria

The field-scale basic reproduction number in the absence of control (Appendix 3) is:

$$R_0 = \text{Basic reproduction number} = R_0^H + R_0^V = \frac{\beta N}{\gamma} + p, \tag{41}$$

with distinct components corresponding to horizontal (R_0^H) and vertical (R_0^V) transmission ([52]; Appendix 3). This applies for both the "strategy vs." and "grower vs." models; below this threshold, there is a disease-free equilibrium (DFE) where there are no infected fields and consequently no controllers (as there is no need to control from disease where there is no risk of infection) (Figure 2(A) - (C)). Other, disease-endemic equilibria are possible and depend on the type of comparison being used.

There are up to four different equilibria, defined as:

- **Disease-free equilibrium** at which the disease is not able to spread even when there is no control via clean seed. This is stable for $R_0 < 1$.
- **Control-free, disease-endemic equilibrium** at which the disease is endemic, but it is unprofitable to use the CSS and so no growers control.
- **All-control, disease-endemic equilibrium** at which all growers control, but nevertheless disease is still endemic in the system.
- **Two-strategy, disease-endemic equilibrium** at which both disease and control equilibrate at some intermediate level.

Which subset of these outcomes are possible will depend on which model formulation is used.

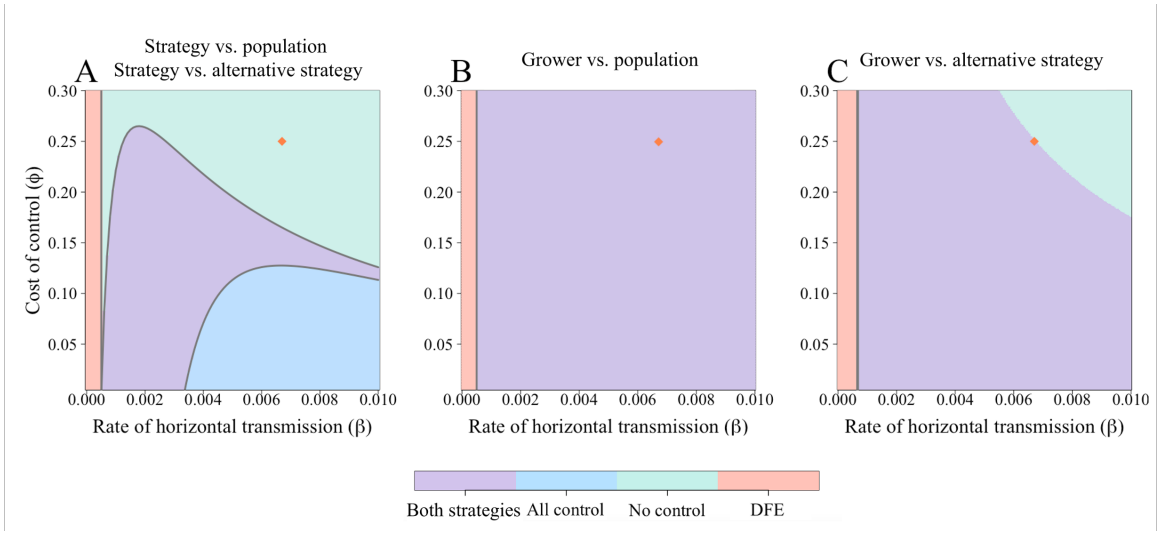

Figure 2: Effect of the rate of horizontal transmission (β) and the cost of control (ϕ) on final equilibrium attained. (A) The "strategy vs. population" and "strategy vs. alternative" models have four possible, mutually-exclusive equilibria. For any given parameter set, both of these models will have the same equilibrium values. Which is attained depends both on R_0 and the balance of benefit between controlling and not controlling. Below $R_0 = 1$ (which corresponds to the vertical line at $\beta = 0.0005$ day $^{-1}$), only the disease-free equilibrium (DFE) is possible. Boundaries between equilibria are determined by the stability conditions (Appendix 3). Also shown in (B) and (C) are the equilibria for the "grower vs." models, though numerical expressions determining these cannot be derived analytically. Parameters are as in Table 1; the orange diamonds mark the default values β and ϕ (which, in (C), is in the "no control" region). The dynamics for these default values are shown in Figure 3. (B) does not show all equilibria achievable for the "grower vs. population" model (see Appendix 3 Figure 1).

Parameterisations leading to each of the four equilibria are possible for the 'strategy vs." models, depending on the balance between the rate of horizontal transmission (β) and the cost of control (ϕ) (Figure 2(A)). In these models, coupling of the switching terms allows mathematical analysis (Appendix 3), and stability conditions can be derived for the single-strategy equilibria. As noted above, although the dynamics in reaching equilibrium can differ, the final state of the system does not depend on whether the "strategy vs. population" or "strategy vs. alternative strategy" model is considered. In each case, for an equilibrium with both strategies present, the expected payoffs (P_C and P_N) must be equal, otherwise one group of growers will continue to be incentivised to change strategy. For only controllers to be present, $P_C > P_N$ and vice versa for the non-controller equilibria. The differences in dynamics approaching equilibrium arise from the different calculations

of Δ (Equations 25 - 28).

Importantly, the "all control" equilibrium, can be attained in these "strategy vs." models. This equilibrium is due to the imperfect protection provided by the control scheme; if horizontal infection is sufficiently significant, even all growers controlling it not guarantee the system remains disease-free. We note here that the equilibrium attained is independent of responsiveness of growers, η, though it does affect the dynamics approaching equilibrium (Appendix 1 Figure 1(D)).

The more complex form of the "grower vs." models makes mathematical analysis difficult, as conditions for model equilibria no longer simply depend on the difference in profits between controllers and non-controllers. However, extensive numerical work (Appendix 3) indicates the equilibrium that is reached again depends only on model parameters (i.e. there is no bistability between equilibria). In the "grower vs." models, only three of the four equilibria can be attained: the "all control" equilibrium is impossible because growers who controlled but nevertheless became infected always consider switching strategy, as they are earning the lowest possible payoff (P_{IC}, the "sucker's payoff"). As long as there is a non-zero probability of infection (which is necessary for control to be worthwhile) there will always be non-controlling growers. Additionally, for the "grower vs. population" model, the "no control" equilibrium is only possible if all non-controllers are infected at equilibrium (Appendix 3). This requires a very narrow range of parameters (Appendix 3 Figure 1), and is not shown by Figure 2(B). Unlike in the "strategy vs." models, where the equilibrium is not dependent on the value of responsiveness of growers, for the "growers vs." models η does affect the final equilibrium values (Appendix 4 Figure 1).

3.1.2 Default behaviour of models.

Using the default parameters, we obtain a "no control" equilibrium when growers employ the "strategy vs. population" and "strategy vs. alternative" models (Figure 3(A)-(B) respectively). At these parameter values, the rate of horizontal transmission and probability of vertical infection are so high that growers are likely to become infected, thus paying the cost of control (ϕ) and the loss due to

disease (L). Therefore, control is not worthwhile.

For the "grower vs. population" model, around 38% of growers control at equilibrium. For these parameters, any non-controlling grower with an infected field (I_N) has a non-zero probability of changing strategy as the presence of uninfected non-controllers (S_N) means that they will earn below the expected population profit, P (see Equation 30). Thus, there will be controllers at equilibrium. Only when the parameter values mean that there will be no susceptible growers will there be a "no control" equilibrium (Equation 56).

This is not observed in the "grower vs. alternative" model, where no growers control at equilibrium (Figure 3(D)). For this model, infected non-controlling growers (I_N) need only earn more than the alternative strategy, so if there is a high number of infected controllers I_N should not switch strategy (Equation 55).

3.2 Effect and implications of epidemiological and economic parameters on uptake of control.

We now continue our investigation using only the "grower vs. alternative" model set-up. The idea of the "reflexive producer" means that growers are likely to respond to previous season's outcomes ([18]), but still incorporate information about what the likely infection risk is for the next season ([53],[20]). We therefore believe that this is a useful approximation of how growers may actually perceive profitability, though we include analysis of the other models in Appendix 4 Figure 3.

3.2.1 Effect of epidemiological and economic parameters on uptake of control.

We investigate the change in uptake of control and proportion of disease fields when the cost of control (ϕ), probability of vertical transmission (p) and loss due to disease (L) are varied (Figure 4). Responses to changes in parameter values were often intuitive. For example, at very high costs of control (ϕ), fewer growers controlled for disease (Figure 4(A)). Similarly, when the risk of vertical transmission increased, so too did the proportion of growers controlling (Figure 4(B)). However, even

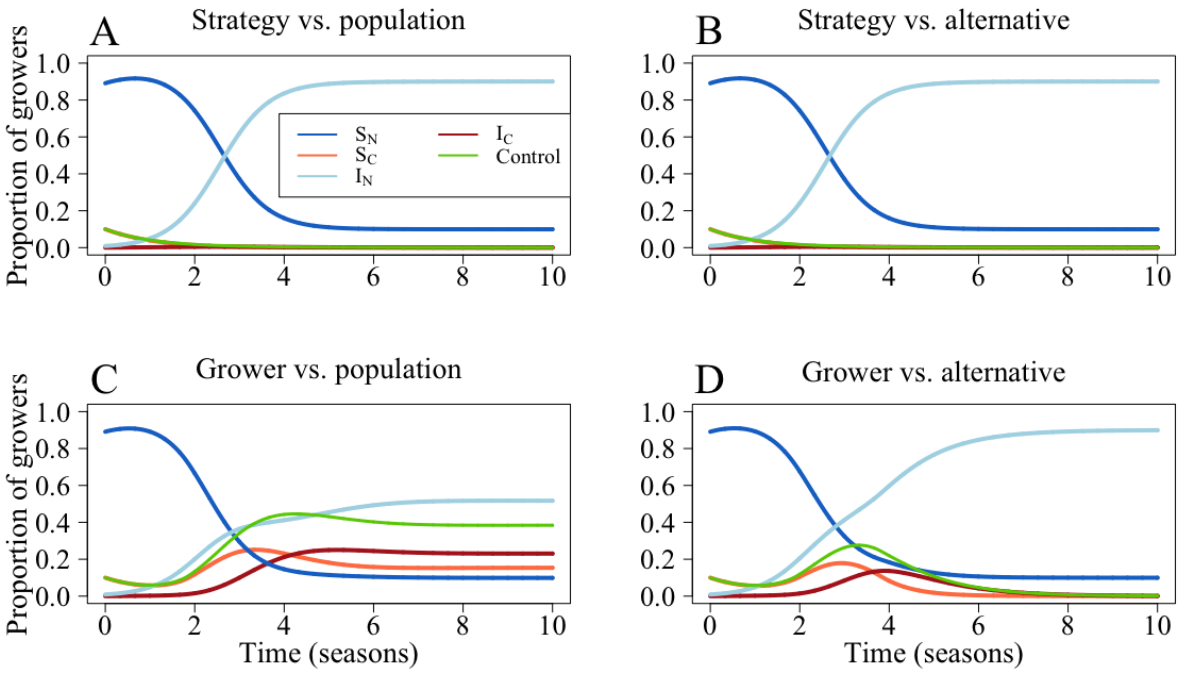

Figure 3: Model dynamics for different means of comparison. Here, S_N are susceptible non-controllers, I_N infected non-controllers, S_C are susceptible controllers and I_C infected controllers. C represents $S_C + I_C$. In (A), growers compare the expected profit of their strategy with that of the entire population; in (B) they compare the expected profit of their strategy with the expected profit of the alternative strategy. For (C), growers compare their own outcome with expected profit of the population and in (D) they compare their outcome with the expected value of the alternative strategy. In (A), (B) and (D), no growers control for disease at equilibrium, though in (C) around 38% of growers control. Parameters are outlined in Table 1.

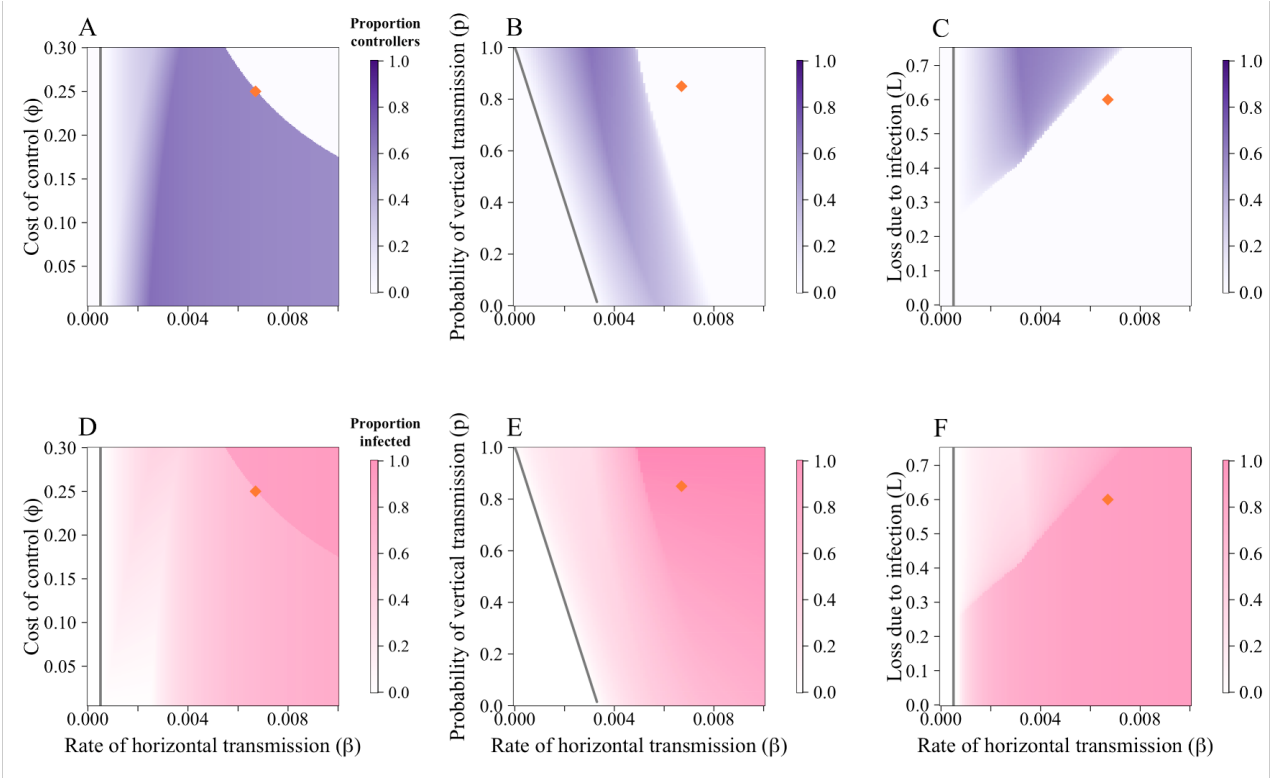

Figure 4: CSS participation and proportion of infected fields are dependent on economic and epidemiological parameters. Plots show the equilibrium proportion of controllers in response to variations in parameters when comparisons are made using the "grower vs. alternative" comparison. (A) Effect of the rate of horizontal transmission (β) and cost of control (ϕ). (B) Effect of the rate of horizontal transmission (β) and probability of receiving infected planting material via trade (p). (C) Effect of the rate of horizontal transmission (β) and loss due to infection (L). At these default parameter values (particularly the high cost of control), both p and L only affects participation in the CSS for a narrow range of β. (D) - (F) show the response of the proportion of infected fields ($(I_N + I_C)/N$) to β, ϕ, p and L respectively. In each case, as control increased, disease decreased. Parameters values are in Table 1; the orange diamonds mark the default values.

within these broad trends, there were also non-monotonic responses. We found that at medium-to-low values for cost of control (ϕ) and the rate of horizontal infection (β), high proportions of growers controlled for disease (Figure 4(A)). Yet as β increased, the higher probability of infection narrows the range of costs for which a controller will consider participation in the CSS, as the grower will likely have to pay both the cost of control (ϕ) and the loss due to disease (L) (the "sucker's payoff").

A similar principle underpins Figure 4(B)-(C). At higher probabilities of horizontal transmission

(β), disease pressure increases and controllers are likely to be infected, even if non-controllers are more likely to be infected due to high probabilities of vertical transmission (p). In each case, when there is a higher proportion of controllers at equilibrium, there are fewer infected fields (Figure 4(D)-(F)).

Growers receive the highest profits when $R_0 < 1$ (i.e. there is no disease, so growers cannot incur any yield loss) (Figure 5). When disease is endemic to the system (i.e. $R_0 > 1$), both profit and yield decrease with higher costs, losses and probability of infection (corresponding to the higher proportion of infected fields in Figure 4(D)-(F)). The lowest yield and profit are obtained when all fields are infected, and at this point no-one controls for disease.

3.2.2 Implications for encouraging adoption of control.

Growers are responsive to both economic and epidemiological factors. If the cost of the control mechanism is too high, growers will not use the scheme as they do not perceive it to be worthwhile. If the scheme provides too little a benefit over the non-control option (which in this case would mean that the probability of vertical transmission was relatively low), growers will also not use it. Though lower losses are beneficial for profit and yield, they discourage disease control. If losses are sufficiently low (such that $L < \phi$), control would not be a rational choice. It is therefore important to emphasise the importance of subsidies for control schemes (which will have the effect of reducing ϕ) to policy-makers in order to encourage participation. The effect of such a subsidy is shown in Appendix 4 Figure 4, where the lower cost of control widens the values of the rate of horizontal transmission (β) and the probability of vertical transmission (p) for which control is profitable.

3.3 Effect of systematic uncertainty

Even when growers underestimate the rate of horizontal transmission, β, to the extent that they do not believe there to be any risk of horizontal transmission (i.e the relative perception of β is zero

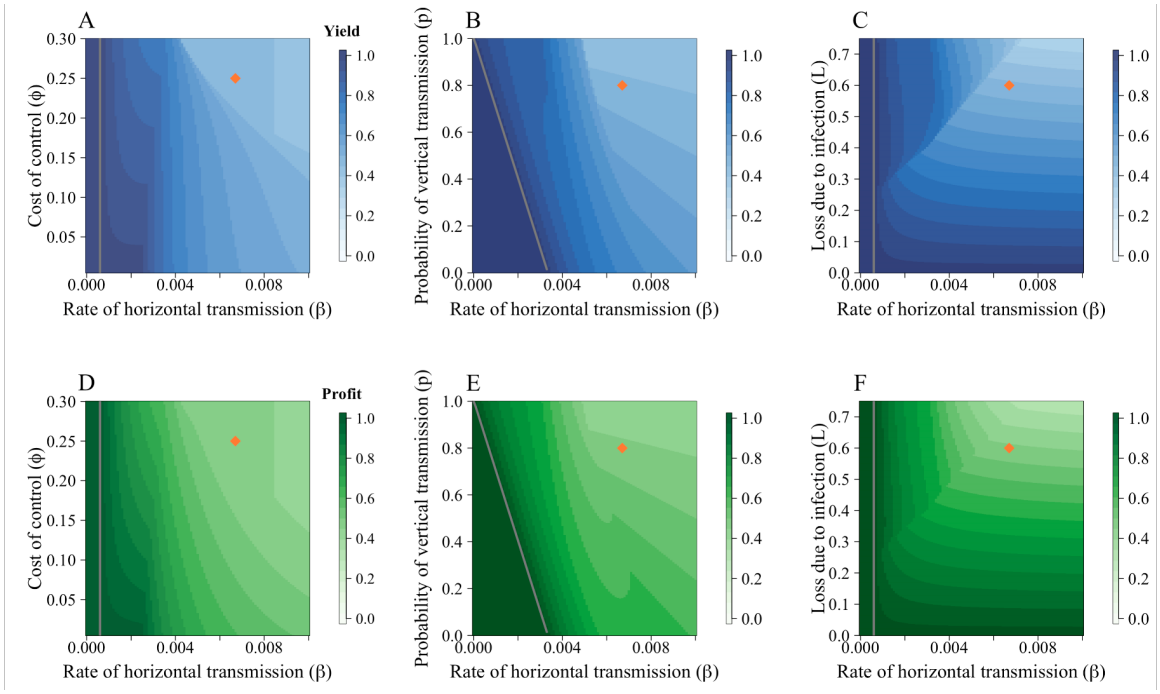

Figure 5: Effect of parameters on yield and profit for the "grower vs. alternative" models. (A) – (C) show the average yield, whilst (D) – (F) show the average profit. For each parameter scan, at low rates of horizontal transmission (β), average yield and profit was higher as very few fields were infected (and thus did not incur the loss due to disease, L). (A) The effect of the cost of control, ϕ, on the yield. As the cost of control increases, fewer growers participate in the CSS and disease incidence increases, leading to lower yields. (B) The effect of the probability of vertical transmission, p. At higher values of p, yield falls as disease prevalence increases. (C) Effect of loss due to disease, L. Intuitively, at higher L, yield falls. Parameters are as in Table 1; the orange diamonds mark the default values. Patterns were similar across (D) – (F), with lower profits observed as ϕ, p and L increased.

($\nu_\beta = 0$, so the perceived value of β (q_β)= 0), some still use the CSS as the probability of vertical transmission is non-zero (Figure 6(A)). As ν_β increases, fewer growers use the control scheme as they perceive they will achieve the "sucker's payoff". The default parameterisation (when $q_\beta = \beta$, i.e. $\nu_\beta = 1$) means that growers should never control. The higher cost of control (ϕ) caused the system to reach a "no control" equilibrium Figure 6(A)).

Unlike when misestimating β, misestimations of I affect perception of both horizontal and vertical transmission. This means there is a less straightforward response to misestimations of I. The kinks in (Figure 6(B)) are caused by changes in the values of the switching terms (Equations 33-36), which set the probability that a grower managing a field of a particular type will switch strategy.

When the relative perception of I, ν_I, is small and increasing, there are more controllers at equilibrium (Figure 6(B)). At low values of ν_I, the non-infected controllers (S_C) growers perceive that the profit for non-controllers is greater than what they have earned ($\tilde{P}_N > P_{SC}$), so they should still consider switching strategy (i.e. $z_{SC} > 0$) (Figure 6(C)). As ν_I increases, their perception of $\tilde{P}_N$ falls and fewer S_C growers switch. This causes an increase in infected controllers (I_C), fewer S_C switch strategy before they become infected. Eventually, $\tilde{P}_N < P_{SC}$ and S_C growers stop switching strategy entirely (around season 4 in Figure 6(D)). Many of these controllers will then become infected and switch strategy, leading to lower participation in control with higher ν_I. As ν_I increases, I_N growers are less likely to switch strategy as $\tilde{P}_C$ falls. Eventually, $z_{IN} = 0$ and no I_N growers should switch strategy (Figure 6(F)). As all I_C growers should have a non-zero probability of switching strategy, but no non-controllers should switch, this leads to a "no control" equilibrium.

4 Discussion

Surprisingly, few plant disease studies examine how human behaviour alters the dynamics of epidemic models. Here we presented a model that uses game theory to incorporate grower decision-making

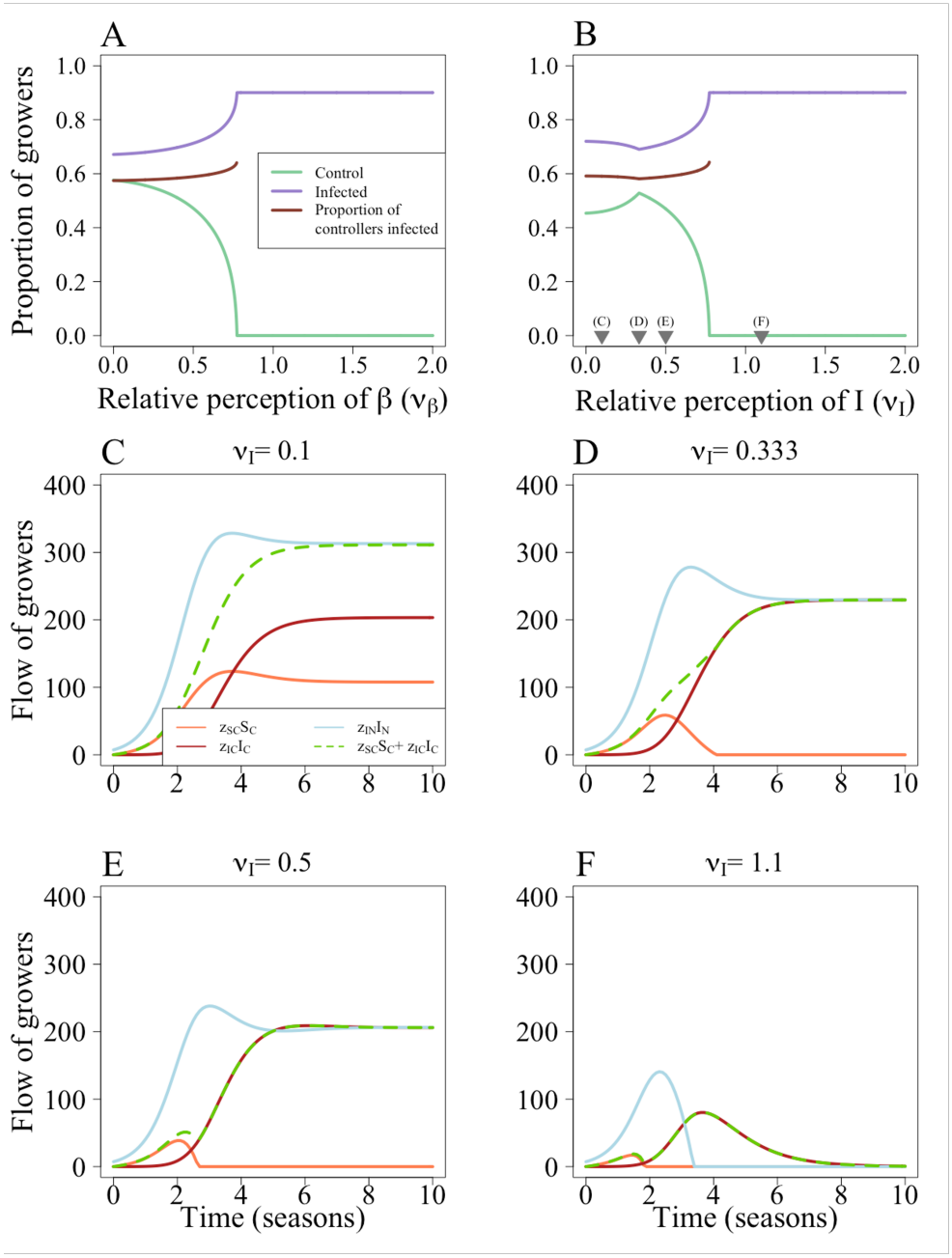

Figure 6: Effect of systematic uncertainty of epidemiological parameters and quantities. When the perception of β (ν_β) is small there is a higher proportion of controllers (A), which decreases as ν_β increases. In (B), there is initially an increase in the proportion controlling with the perception of I (ν_I). In (B) the number of infected fields levels out once the conditions for a "no control" equilibrium are met. The grey triangles in (B) show the values of ν_I used in (C)-(F). (C) When $\nu_I = 0.1$, S_C growers still perceive that it is profitable to change strategy ($z_{SC} > 0$). (D) At $\nu_I = 0.33$, as the number of infected fields increases, the perceived value of P_N ($\tilde{P}_N$) decreases and $z_{SC} = 0$. However, they continue to get infected and there is an increase in I_C fields. (E) At $\nu_I = 0.5$, S_C fields stop switching strategy earlier in the epidemic. As $\tilde{P}_C$ falls, so does z_{IN}. (F) At $\nu_I = 1.1$, $\tilde{P}_C < \tilde{P}_{IN}$, so $z_{I_N} = 0$. This leads to a "no control" equilibrium. Parameters are as in Table 1

into a model of CBSD and investigated the effect of economically- and epidemiologically- important parameters on disease spread and the uptake of control by way of a clean seed system. We found that the formulation of the model (whether based on rational or strategic-adaptive expectations) had a significant consequences on the possible equilibria that could be attained. The model parameters also influenced outcomes, and interactions between them (particularly the cost of control and the rate of horizontal transmission) had large impacts on growers' participation in the CSS. Finally, the response to misperception of epidemiological quantities was often not intuitive, and could have implications for the communication of risk.

We found that basing decisions on rational (the "strategy vs." models) or strategic-adaptive ("grower vs." models) changed the potential outcomes for the model. Only the "strategy vs." comparisons permitted an "all control" equilibrium, whereas the low payoff for controllers with infected fields (the "sucker's payoff") meant that they will always have a non-zero probability of switching strategy in the "grower vs." comparisons. This difference in possible equilibria from the models shows that it is important not only to include behaviour in epidemic models, but to consider how it is included based on what quantities are being compared and what information will be available to individual growers, such as the disease pressure, profit of other growers and breakdown of the expected profit for each strategy.

Game theory predicts that high participation in voluntary disease control schemes is difficult to attain, as the success of disease control policies in reducing the probability of infection disincentiveses continued use of the scheme ([22], [54], [25]; reviewed in [7]). Indeed, in the case of voluntary vaccinations, at least with perfect immunity, models predict that actual level of vaccination will be less than that needed for herd immunity [2]. This is because of the positive externalities generated by having high levels of vaccination; if the majority of individuals are perfectly protected from infection, then there will be a lower probability of infection for others who are not vaccinated. They can then get much of the benefit of vaccination without themselves incurring any risk or cost. Indeed, the risk of other individuals "free-riding" off an individual's control mechanism may also be a disincentive

to control ([53]).

Yet if vaccines are not completely effective, the reduction in magnitude of positive externalities can lead to increased vaccine uptake until a stable "all vaccinated" equilibrium is attained [55]. That is, as the vaccine no longer confers perfect immunity to the vaccinated, they may then transmit the disease to the unvaccinated. This decreases the benefit of high levels of vaccination in the population to the unvaccinated, and thus encourages them to be engaged in disease prevention themselves. This is more akin to the control system we have considered here, where the CSS only reduces the initial probability of infection and does not provide continued control during the season.

In our model, it was the imperfect protection offered by the CSS that allowed the "all control" equilibrium to be attained in the "strategy vs." models (Figure 2(A)). Although the CSS eliminates a grower's chance of receiving infected material via vertical transmission, it offers no protection from infection via the whitefly vector. By comparing the expected payoff of everyone using their strategy, rather than basing decisions on their own outcomes, even growers who themselves received the lowest possible payoff do not necessarily have a positive probability of switching strategy.

In the "strategy vs." models, we found that use of CSS could reduce CBSD spread, but never lead to disease elimination. By considering the stability of the equilibria, we can see that for an "all control" equilibrium, disease must be capable of moving via the whitefly vector alone. Fields that were initially planted with clean seed can therefore become infected, allowing disease to remain endemic in the system. Thus, though increased use of the CSS leads to lower disease prevalence, some degree of disease spread is required for growers to consider the CSS worthwhile. The threshold that determines the benefit of the CSS is determined by the stability conditions outlined in Appendix 3, which are broadly determined by the probability of incurring the loss due to disease for both controllers and non-controllers. This result, and associated requirement for horizontal transmission, holds true for the "grower vs." models, too, though analytical expressions cannot be derived to describe the stability of the equilibria.

We assumed that all growers had the same responsiveness (our parameter η). This impacted

the final equilibrium value attained in our "grower vs." models (Appendix 4 Figure 1), but not our "strategy vs." models (Appendix 1 Figure 1). We have also assumed that growers would always choose the most profitable strategy; if the alternative strategy has a higher expected profit, growers have a non-zero probability of switching (Figure 1(C)). As η increases, growers are almost guaranteed to switch strategy. For the "strategy vs. models", which are based on rational expectations, this increase in η approaches the assumption of perfect rationality in game theory [24]. Both rationality and responsiveness are likely to vary from grower to grower. Their distribution can be described by a "risk attitude" framework, in which individuals fall on a spectrum between risk tolerant and risk averse [56]. Including such heterogeneity in risk perception in future studies would allow for greater understanding of behaviour in uncertain environments.

When conducting our further analysis into the effects of parameters and misperception, we focused on the "grower vs. alternative" evaluation, where individual growers compare their own profits with the expected profit of the alternative strategy. The reasoning behind this was twofold: first, this has been used previously in plant epidemiological literature ([12], [13] and [17]). Secondly, and most importantly, we believe this comparison to be most similar to what real growers might be expected to do ([18], [53], [20]). It requires a moderate amount of information, as it only relies on the individual knowing their own profit from the previous season and the expected profit of the alternative strategy (Table 2).

Epidemiological and economic parameters had a substantial impact on the uptake of control (Figure 4). When the cost of control (ϕ) was low, more growers used the CSS (Figure 4(A) and Appendix 4 Figure 4). However, even at these low costs of control, if the rate of horizontal transmission was too high, growers would still not use the control scheme as the high probability of infection meant they were likely to incur the dual penalty of yield loss and cost of control. Subsidies, then are likely to be an effective way of encouraging control, particularly as both profits and yields are highest when the cost of control is low (Figure 5(A) and (C)) In practice, the success of subsidies may be limited by a lack of information amongst growers, unwillingness of growers to vary from current

practices, or perceived ineffectiveness of control schemes. It is important that the introduction of subsidies control is accompanied by outreach and educational campaigns for growers ([57]).

For control to be worthwhile, it must provide sufficient benefit compared to the alternative non-control strategy. In our case study, the benefit of control is the elimination of vertical transmission for those that use the CSS. Thus, if the probability of vertical transmission is sufficiently low for non-controllers, there may not be enough incentive to control (Figure 4(B)). Similarly, if growers do not stand to lose much yield when infected, the cost of control may be a disincentive, particularly if the probability of infection is low (Figure 4(C)). Indeed, if the expected yield loss over a season is less than the cost of control, it would never be worthwhile to control. It is therefore important to emphasise the benefits of control to growers, particularly if there is likely to be a significant future increase in disease prevalence that they are not accounting for when they are making their decisions.

Growers' inability to properly perceive epidemiological parameters had varying effects on CSS participation. Underestimating the rate of horizontal transmission, β, increased participation as growers believed that infection via this route was unlikely (Figure 6(A)). However, there was still a high probability of vertical transmission so control was worthwhile. As estimates of β increased, control was less favourable and participation decreased. A similar trend was observed in the when estimating α, the whitefly's dispersal scale parameter in a more complex spatial-stochastic formulation of our model (Appendix 2 Figure 3). Misestimations of the proportions of infected individuals had a non-monotonic pattern, with the proportion of controllers initially increasing with the perceived number of infected fields ($q_I I$) before falling as more controllers become infected and thus receive the lowest payoff (Figure 6(B)).

It is therefore important to understand how growers perceive their probability of infection, as this will influence their participation in control schemes and subsequent spread of disease. Overestimating (or exaggerating when communicating with growers) the threat can potentially discourage participation due to the perceived inevitability of infection regardless of any actions taken to prevent it. The effect of underestimation depends on the route of disease transmission in question,

but control only remains attractive if it is seen as providing sufficient benefits. Again, here every grower has the same level of misperception, but in reality this will be closely linked to an individual grower's attitude towards risk, access to information and previous history of infection, which all will be highly variable between growers.

We have only considered economic incentives for grower participation in the CSS. Realistically, other considerations - such as preference for local varieties, flavour, *etc.* - will play an important role in determining participation. Indeed, surveys of cassava growers in Sub-Saharan Africa have found that such qualitative traits have a greater role in determining a grower's preference than economic traits such as yield [14], [58]. These private preferences or perceptions of risk are difficult to assess and could be encapsulated in a "cost of control" parameter that will vary from grower to grower.

As well as economic considerations, there are many social factors underpinning the adoption of a new control mechanism. Descriptive norms (what the majority of people are doing, and thus what is deemed socially acceptable; [59]) have been found to influence growers' decision to partake in control ([60]). Relatedly, a grower's social network and ability to access trusted information has a strong impact on management practices. Some growers will treat the advice of "experts" with a degree of scepticism (often due to a perceived conflict of interest or disconnectedness from the "true" needs of growers) ([19]). Other grower have developed a strong dependency on the advice of these experts ([19]). Such contrasting beliefs and degree of trust in knowledge sources, even in adjacent landscapes, highlights the heterogeneous information network used by growers. In all cases, growers placed high value on personal and trans-generational knowledge [19]. Growers may also be reluctant to adopt a new technology for fear of failure, or as they contradict a pre-existing ideology [19].

Kaup introduces the idea of a "reflexive producer" as a grower that must make decisions by balancing knowledge generated by local growers and knowledge provided by external experts ([18]). This is similar to the hypothetical scenarios we previously outlined, particularly that of the "grower vs. alternative" scenario where a grower is using a particular control strategy and is then approached by an extension worker that proposes an different strategy. Thus, motivations and considerations of

a grower to partake in specific control schemes are highly varied and depend on the market demands, information network available to growers, disease pressure and proximity to previous outbreaks [20].

For simplicity, in both our stochastic and deterministic models we included an average loss of yield. In reality, yield will depend on how far into the growing season the field was infected and the rate of within-field disease spread (as well as many other factors, such as environmental stochasticity, the variety of cassava in question, availability of technology, planting density *etc.* ([61],[62], [63])). The inclusion of time-dependent yield loss will favour clean seed use, as these fields will on average be infected later in the growing season than conventional fields. Indeed, this could shift the balance of payoffs such that the payoff for an infected field planted with clean seed is no longer the lowest possible payoff. If this shift were observed, an "all control" equilibrium may now be stable in the "grower vs." models that we have described.

Though we have not investigated it here, CSS can "bundle" desirable traits by providing CBSD resistant or tolerant planting material, changing the nature of the externalities generated by clean seed use. There are few such varieties available for cassava. Provision of resistant material, though of obvious benefit to the recipient, will also increase the benefit experienced by free riders by reducing their probability of infection. Cassava varieties that are tolerant to CBSD infection will benefit the focal grower, but may increase the disease pressure experienced by other growers and thus incentivise control. These are potentially important distinctions, which we intend to return to in future work.

The inclusion of grower behaviour alters disease dynamics, as the option to participate in a control scheme lowered the probability of infection and has the potential to increase the yield achieved by the growers. We have found that the means of inclusion of behaviour is also important; the quantities of comparison impact dynamics approaching equilibrium, but also determine the nature of the equilibria that can be achieved. Our investigation has been limited to one case study, but the decision modelling framework laid out in this paper is sufficiently flexible to be incorporated into a range of other disease systems. It would also allow grower behaviour to feed into other types of study, for example those in which control has a spatial element and/or is done reactively in response

to detected infection ([64], [65], [66], [67]), or where estimates of parameters become more accurate over time and as disease spreads [68].

5 Acknowledgements

REMW acknowledges the Biotechnology and Biological Sciences Research Council of the United Kingdom (BBSRC; `https://bbsrc.ukri.org/`) for support via a University of Cambridge DTP PhD studentship (Project Reference 2119272). All authors would like to thank Neil McRoberts and two anonymous reviewers for their helpful comments that improved the manuscript.

References

[1] Bauch CT, Galvani AP, Earn DJ. Group interest versus self-interest in smallpox vaccination policy. Proceedings of the National Academy of Sciences. 2003;100(18):10564-7.

[2] Bauch CT, Earn DJD. Vaccination and the theory of games. Proceedings of the National Academy of Sciences. 2004;101(36):13391-4.

[3] Salathé M, Bonhoeffer S. The effect of opinion clustering on disease outbreaks. Journal of The Royal Society Interface. 2008;5(29):1505-8.

[4] d'Onofrio A, Manfredi P, Poletti P. The impact of vaccine side effects on the natural history of immunization programmes: an imitation-game approach. Journal of theoretical biology. 2011;273(1):63-71.

[5] Cascante-Vega J, Torres-Florez S, Cordovez J, Santos-Vega M. How disease risk awareness modulates transmission: coupling infectious disease models with behavioural dynamics. Royal Society open science. 2022;9(1):210803.

[6] Funk S, Salathé M, Jansen VA. Modelling the influence of human behaviour on the spread of infectious diseases: a review. Journal of the Royal Society Interface. 2010;7(50):1247-56.

[7] Chang SL, Piraveenan M, Pattison P, Prokopenko M. Game theoretic modelling of infectious disease dynamics and intervention methods: a review. Journal of Biological Dynamics. 2020;14(1):57-89. PMID: 31996099. Available from: `https://doi.org/10.1080/17513758.2020.1720322`.

[8] Hidano A, Enticott G, Christley RM, Gates MC. Modeling Dynamic Human Behavioral Changes in Animal Disease Models: Challenges and Opportunities for Addressing Bias. Frontiers in Veterinary Science. 2018;5. Available from: https://www.frontiersin.org/article/10.3389/fvets.2018.00137.

[9] Bucini G, Merrill SC, Clark E, Moegenburg SM, Zia A, Koliba CJ, et al. Risk attitudes affect livestock biosecurity decisions with ramifications for disease control in a simulated production system. Frontiers in Veterinary Science. 2019:196.

[10] Delabouglise A, Boni MF. Game theory of vaccination and depopulation for managing livestock diseases and zoonoses on small-scale farms. Epidemics. 2020;30:100370.

[11] Mohr S, Beard R, Nisbet AJ, Burgess STG, Reeve R, Denwood M, et al. Uptake of Diagnostic Tests by Livestock Farmers: A Stochastic Game Theory Approach. Frontiers in Veterinary Science. 2020;7:36. Available from: https://www.frontiersin.org/article/10.3389/fvets.2020.00036.

[12] Milne AE, Bell JR, Hutchison WD, van den Bosch F, Mitchell PD, Crowder D, et al. The Effect of Farmers' Decisions on Pest Control with Bt Crops: A Billion Dollar Game of Strategy. PLOS Computational Biology. 2016 12;11(12):1-18.

[13] McQuaid CF, Gilligan CA, van den Bosch F. Considering behaviour to ensure the success of a disease control strategy. Royal Society Open Science. 2017a;4(12):170721. Available from: https://royalsocietypublishing.org/doi/abs/10.1098/rsos.170721.

[14] Szyniszewska AM, Chikoti PC, Tembo M, Mulenga R, Gilligan CA, van den Bosch F, et al. Cassava planting material movement and grower behaviour in Zambia: implications for disease management. bioRxiv. 2019. Available from: https://www.biorxiv.org/content/early/2019/03/06/528851.

[15] Milne AE, Gottwald T, Parnell SR, Alonso Chavez V, van den Bosch F. What makes or breaks a campaign to stop an invading plant pathogen? PLOS Computational Biology. 2020 02;16(2):1-20. Available from: https://doi.org/10.1371/journal.pcbi.1007570.

[16] Bate AM, Jones G, Kleczkowski A, Touza J. Modelling the effectiveness of collaborative schemes for disease and pest outbreak prevention. Ecological Modelling. 2021;442:109411. Available from: https://www.sciencedirect.com/science/article/pii/S0304380020304701.

[17] Saikai Y, Hurley TM, Mitchell PD. An agent-based model of insect resistance management and mitigation for Bt maize: a social science perspective. Pest Management Science. 2021;77(1):273-84. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1002/ps.6016.

[18] Kaup BZ. The Reflexive Producer: The Influence of Farmer Knowledge Upon the Use of Bt Corn. Rural Sociology. 2008;73(1):62-81. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1526/003601108783575871.

[19] Sherman J, Gent D. Concepts of Sustainability, Motivations for Pest Management Approaches, and Implications for Communicating Change. Plant Disease. 2014 08;98:1024-35.

[20] Garcia-Figuera S, Deniston-Sheets H, Grafton-Cardwell E, Babcock B, Lubell M, McRoberts N. Perceived Vulnerability and Propensity to Adopt Best Management Practices for Huanglongbing Disease of Citrus in California. Phytopathology. 2021;111(10):1758-73. PMID: 33599529. Available from: https://doi.org/10.1094/PHYTO-12-20-0544-R.

[21] Siegal G, Siegal N, Bonnie RJ. An Account of Collective Actions in Public Health. American Journal of Public Health. 2009;99(9):1583-7. PMID: 19608946. Available from: https://doi.org/10.2105/AJPH.2008.152629.

[22] Geoffard PY, Philipson T. Disease Eradication: Private versus Public Vaccination. The American Economic Review. 1997;87(1):222-30. Available from: http://www.jstor.org/stable/2950864.

[23] Singerman A, Useche P. The Role of Strategic Uncertainty in Area-wide Pest Management Decisions of Florida Citrus Growers. American Journal of Agricultural Economics. 2019;101(4):991-1011. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1093/ajae/aaz006.

[24] Morris P. Introduction to Game Theory. Universitext. Springer; 2012.

[25] Bauch CT. Imitation dynamics predict vaccinating behaviour. Proceedings of the Royal Society B: Biological Sciences. 2005;272(1573):1669-75.

[26] Riechert SE, Hammerstein P. Game Theory in the Ecological Context. Annual Review of Ecology and Systematics. 1983;14(1):377-409. Available from: https://doi.org/10.1146/annurev.es.14.110183.002113.

[27] Roca CP, Cuesta JA, Sánchez A. Evolutionary game theory: Temporal and spatial effects beyond replicator dynamics. Physics of Life Reviews. 2009;6(4):208-49. Available from: https://www.sciencedirect.com/science/article/pii/S1571064509000256.

[28] Hummert S, Bohl K, Basanta D, Deutsch A, Werner S, Theißen G, et al. Evolutionary game theory: cells as players. Molecular BioSystems. 2014;10(12):3044-65.

[29] Anderson M. The evolution of eusociality. Annual Review of Ecology and Systematics. 1984;15(1):165-89.

[30] Nowak MA, Bonhoeffer S, May RM. Spatial games and the maintenance of cooperation. Proceedings of the National Academy of Sciences. 1994;91(11):4877-81. Available from: https://www.pnas.org/doi/abs/10.1073/pnas.91.11.4877.

[31] Pacheco JM, Santos FC, Dingli D. The ecology of cancer from an evolutionary game theory perspective. Interface focus. 2014;4(4):20140019.

[32] Mlambo L. Adaptive and Rational Expectations Hypotheses: Reviewing the Critiques. Journal of Economic Behaviour. 2012 01.

[33] Fenichel EP, Wang X. In: Manfredi P, D'Onofrio A, editors. The Mechanism and Phenomena of Adaptive Human Behavior During an Epidemic and the Role of Information. New York, NY: Springer New York; 2013. p. 153-68. Available from: https://doi.org/10.1007/978-1-4614-5474-8_10.

[34] Council NR, et al. Strategic planning for the Florida citrus industry: addressing citrus greening disease. National Academies Press; 2010.

[35] Florec V, Sadler RJ, White B, Dominiak BC. Choosing the battles: The economics of area wide pest management for Queensland fruit fly. Food Policy. 2013;38:203-13. Available from: https://www.sciencedirect.com/science/article/pii/S0306919212001212.

[36] Monger WA, Seal S, Isaac AM, Foster GD. Molecular characterization of the Cassava brown streak virus coat protein. Plant Pathology. 2001;50(4):527-34.

[37] Patil BL, Ogwok E, Wagaba H, Mohammed IU, Yadav JS, Bagewadi B, et al. RNAi-mediated resistance to diverse isolates belonging to two virus species involved in Cassava brown streak disease. Molecular Plant Pathology. 2011;12(1):31-41.

[38] Hillocks RJ, Raya MD, Mtunda K, Kiozia H. Effects of Brown Streak Virus Disease on Yield and Quality of Cassava in Tanzania. Journal of Phytopathology. 2001;149(7-8):389-94. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1439-0434.2001.tb03868.x.

[39] Legg J, Somado E, Barker I, Beach L, Ceballos H, Cuellar W, et al. A global alliance declaring war on cassava viruses in Africa. Food Security. 2014;6:231–248. Available from: https://doi.org/10.1007/s12571-014-0340-x.

[40] Maruthi MN, Hillocks RJ, Mtunda K, Raya MD, Muhanna M, Kiozia H, et al. Transmission of Cassava brown streak virus by Bemisia tabaci (Gennadius). Journal of Phytopathology. 2005;153(5):307-12. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1439-0434.2005.00974.x.

[41] Chipeta MM, Shanahan P, Melis R, Sibiya J, Benesi IRM. Farmers' knowledge of cassava brown streak disease and its management in Malawi. International Journal of Pest Management. 2016;62(3):175-84. Available from: https://doi.org/10.1080/09670874.2016.1167268.

[42] McQuaid CF, van den Bosch F, Szyniszewska A, Alicai T, Pariyo A, Chikoti PC, et al. Spatial dynamics and control of a crop pathogen with mixed-mode transmission. PLOS Computational Biology. 2017b 07;13(7):1-18. Available from: https://doi.org/10.1371/journal.pcbi.1005654.

[43] Legg J, Ndalahwa M, Yabeja J, Ndyetabula I, Bouwmeester H, Shirima R, et al. Community phytosanitation to manage cassava brown streak disease. Virus Research. 2017;241:236-53. Plant Virus Epidemiology. Available from: https://www.sciencedirect.com/science/article/pii/S0168170217301533.

[44] Madden LV, Hughes G, van den Bosch F. The study of plant disease epidemics. St Paul, MN: American Phytopathological Society (APS Press); 2007.

[45] Szyniszewska AM. CassavaMap, a fine-resolution disaggregation of cassava production and harvested area in Africa in 2014. Scientific data. 2020;7:159.

[46] Jeger MJ, Holt J, van den Bosch F, Madden LV. Epidemiology of insect-transmitted plant viruses: modelling disease dynamics and control interventions. Physiological Entomology. 2004;29:291-304.

[47] Ephraim N, Yona B, Evans A, Sharon A, Titus A. Effect of cassava brown streak disease (CBSD) on cassava (Manihot esculenta Crantz) root storage components, starch quantities and starch quality properties. International Journal of Plant Physiology and Biochemistry. 2015.

[48] van den Bosch F, Jeger MJ, Gilligan CA. Disease control and its selection for damaging plant virus strains in vegetatively propagated staple food crops; a theoretical assessment. Proceedings of the Royal Society B: Biological Sciences. 2007;274(1606):11-8. Available from: https://royalsocietypublishing.org/doi/abs/10.1098/rspb.2006.3715.

[49] Gibson RW, Otim-Nape GW. Factors determining recovery and reversion in mosaic-affected African cassava mosaic virus resistant cassava. Annals of Applied Biology. 1997;131(2):259-71. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1744-7348.1997.tb05155.x.

[50] Mohammed IU, Ghosh S, Maruthi MN. Host and virus effects on reversion in cassava affected by cassava brown streak disease. Plant Pathology. 2016;65(4):593-600. Available from: https://bsppjournals.onlinelibrary.wiley.com/doi/abs/10.1111/ppa.12458.

[51] Fu F, Rosenbloom DI, Wang L, Nowak MA. Imitation dynamics of vaccination behaviour on social networks. Proceedings of the Royal Society B: Biological Sciences. 2011;278(1702):42-9. Available from: https://royalsocietypublishing.org/doi/abs/10.1098/rspb.2010.1107.

[52] Hamelin FM, Bowen B, Bernhard P, Bokil VA. Optimal Control of Plant Disease Epidemics with Clean Seed Usage. Bulletin of mathematical biology. 2021;83 5:46.

[53] Milne AE, Teiken C, Deledalle F, van den Bosch F, Gottwald T, McRoberts N. Growers' risk perception and trust in control options for huanglongbing citrus-disease in Florida and California. Crop Protection. 2018;114:177-86. Available from: https://www.sciencedirect.com/science/article/pii/S0261219418302540.

[54] Bauch CT, Galvani AP, Earn DJD. Group interest versus self-interest in smallpox vaccination policy. Proceedings of the National Academy of Sciences. 2003;100(18):10564-7. Available from: https://www.pnas.org/content/100/18/10564.

[55] Wu B, Fu F, Wang L. Imperfect Vaccine Aggravates the Long-Standing Dilemma of Voluntary Vaccination. PLOS ONE. 2011 06;6(6):1-7. Available from: https://doi.org/10.1371/journal.pone.0020577.

[56] Zhang R, Brennan TJ, Lo AW. The origin of risk aversion. Proceedings of the National Academy of Sciences. 2014. Available from: https://www.pnas.org/content/early/2014/11/26/1406755111.

[57] Jowett K, Milne AE, Potts SG, Senapathi D, Storkey J. Communicating Carabids: Engaging farmers to encourage uptake of Integrated Pest Management. Pest Management Science. 2022;n/a(n/a). Available from: https://onlinelibrary.wiley.com/doi/abs/10.1002/ps.6878.

[58] Houngue JA, Pita JS, Cacaï GHT, Zandjanakou-Tachin M, Abidjo EAE, Ahanhanzo C. Survey of farmers' knowledge of cassava mosaic disease and their preferences for cassava cultivars in three agro-ecological zones in Benin. Ethnobiology Ethnomedicine. 2019;14(29). Available from: https://doi.org/10.1186/s13002-018-0228-5.

[59] Lazić A, Kalinova KN, Packer J, Pae R, Petrović MB, Popović D, et al. Social nudges for vaccination: How communicating herd behaviour influences vaccination intentions. British Journal of Health Psychology. 2021;26(4):1219-37. Available from: https://bpspsychub.onlinelibrary.wiley.com/doi/abs/10.1111/bjhp.12556.

[60] Rodriguez JM, Molnar JJ, Fazio RA, Sydnor E, Lowe MJ. Barriers to adoption of sustainable agriculture practices: Change agent perspectives. Renewable Agriculture and Food Systems. 2009;24(1):60–71.

[61] Bull SE, Ndunguru J, Gruissem W, Beeching J, Vanderschuren H. Cassava: Constraints to production and the transfer of biotechnology to African laboratories. Plant cell reports. 2011 05;30:779-87.

[62] Siloko IU, Ukhurebor KE, Siloko EA, Enoyoze E, Bobadoye AO, Ishiekwene CC, et al. Effects of some meteorological variables on cassava production in Edo State, Nigeria via density estimation. Scientific African. 2021;13:e00852. Available from: https://www.sciencedirect.com/science/article/pii/S2468227621001563.

[63] Tafesse A, Mena B, Belay A, Aynekulu E, Recha JW, Osano PM, et al. Cassava Production Efficiency in Southern Ethiopia: The Parametric Model Analysis. Frontiers in Sustainable Food Systems. 2021;5. Available from: https://www.frontiersin.org/article/10.3389/fsufs.2021.758951.

[64] Cunniffe NJ, Stutt ROJH, DeSimone RE, Gottwald TR, Gilligan CA. Optimising and Communicating Options for the Control of Invasive Plant Disease When There Is Epidemiological Uncertainty. PLOS Computational Biology. 2015 04;11(4):1-24. Available from: https://doi.org/10.1371/journal.pcbi.1004211.

[65] Cunniffe NJ, Cobb RC, Meentemeyer RK, Rizzo DM, Gilligan CA. Modeling when, where, and how to manage a forest epidemic, motivated by sudden oak death in California. Proceedings of the National Academy of Sciences. 2016;113(20):5640-5. Available from: https://www.pnas.org/content/113/20/5640.

[66] Hyatt-Twynam SR, Parnell S, Stutt ROJH, Gottwald TR, Gilligan CA, Cunniffe NJ. Risk-based management of invading plant disease. New Phytologist. 2017;214(3):1317-29. Available from: https://nph.onlinelibrary.wiley.com/doi/abs/10.1111/nph.14488.

[67] Laranjeira FF, Silva SXB, Murray-Watson RE, Soares ACF, Santos-Filho HP, Cunniffe NJ. Spatiotemporal dynamics and modelling support the case for area-wide management of citrus greasy spot in a Brazilian smallholder farming region. Plant Pathology. 2020;69(3):467-83. Available from: https://bsppjournals.onlinelibrary.wiley.com/doi/abs/10.1111/ppa.13146.

[68] Thompson RN, Gilligan CA, Cunniffe NJ. Control fast or control smart: When should invading pathogens be controlled? PLOS Computational Biology. 2018 02;14(2):1-21. Available from: https://doi.org/10.1371/journal.pcbi.1006014.

[69] Cunniffe NJ, Gilligan CA. Use of Mathematical Models to Predict Epidemics and to Optimize Disease Detection and Management. In: Ristaino JB, Records A, editors. Emerging Plant Diseases and Global Food Security. APS Press; 2020. .

[70] Fabre F, Coville J, Cunniffe NJ. Optimising reactive disease management using spatially explicit models at the landscape scale. In: Scott P, Strange R, Korsten L, Gullino ML, editors. Plant Diseases and Food Security in the 21st Century:. Springer; 2021. .

[71] Byrne DN. Migration and dispersal by the sweet potato whitefly, Bemisia tabaci. Agricultural and Forest Meteorology. 1999;97(4):309-16. Available from: https://www.sciencedirect.com/science/article/pii/S016819239900074X.

[72] Keeling MJ, Rohani P. Modeling Infectious Diseases in Humans and Animals. Princeton University Press; 2008. Available from: http://www.jstor.org/stable/j.ctvcm4gk0.

[73] Grimmett G, Welsh D. Probability: an introduction. Oxford University Press; 2014.

[74] Byrne DN, Rathman RJ, Orum TV, Palumbo JC. Localized migration and dispersal by the sweet potato whitefly, Bemisia tabaci. Oecologia. 1996;105:320-8.

[75] Shaw MW. Simulation of population expansion and spatial pattern when individual dispersal distributions do not decline exponentially with distance. Proceedings of the Royal Society of London Series B: Biological Sciences. 1995;259(1356):243-8. Available from: https://royalsocietypublishing.org/doi/abs/10.1098/rspb.1995.0036.

[76] Ntawuruhunga P, Legg J, Okidi J, Okao-Okuja G, Tadu G, Remington T. Southern Sudan, Equatoria Region, Cassava Baseline Survey Technical Report. IITA. 2007 08.

[77] Chikoti P, Melis R, Shanahan P. Farmer's Perception of Cassava Mosaic Disease, Preferences and Constraints in Lupaula Province of Zambia. American Journal of Plant Sciences. 2016:1129-38.

[78] Djaha KE, Abo K, Kone T, Kone D, Kone M. Analysis of the population structure of cassava growers, production systems, and plots' sanitary state in Côte d'Ivoire. Journal of Animal & Plant Sciences. 2018:5833-43.

[79] Houngue JA, Pita JS, Cacaï GHT, Zandjanakou-Tachin M, Abidjo EAE, Ahanhanzo C. Survey of farmers' knowledge of cassava mosaic disease and their preferences for cassava cultivars in three agro-ecological zones in Benin. Journal of Ethnobiology Ethnomedicine. 2018.

[80] Teeken B, Olaosebikan O, Haleegoah JAS, Oladejo E, Madu TU, Bello A, et al. Cassava Trait Preferences of Men and Women Farmers in Nigeria: Implications for Breeding. Economic Botany. 2018;72:263 277.

[81] Diekmann O, Heesterbeek JAP, Roberts MG. The construction of next-generation matrices for compartmental epidemic models. Journal of The Royal Society Interface. 2010;7(47):873-85. Available from: https://royalsocietypublishing.org/doi/abs/10.1098/rsif.2009.0386.

[82] van den Driessche P. Reproduction numbers of infectious disease models. Infectious Disease Modelling. 2017;2(3):288-303. Available from: https://www.sciencedirect.com/science/article/pii/S2468042717300209.

A Appendix 1: Underlying behaviour of "strategy vs." models.

A.1 Probability of infection, responsiveness, profits, and switching terms in the "strategy vs." models

To demonstrate the underlying drivers of dynamics in our models, we dissect the behaviour of the "strategy vs. population" model in some depth. We first investigated the response of the probability of infection, the expected profits and the probability of switching strategy to the total number of infected fields (Figure 1). Generally, as the number of infected fields increases, so too does the probability of infection for controllers and non-controllers (q_C and q_N respectively) (Figure 1(A)). However, the increased number of infected fields changes the relative importance of the infection pathways for non-controllers. With an increase in the number of infected fields, the instantaneous probability of horizontal transmission for non-controllers (p_N(Horiz), Equation 19), decreases as field infection has already taken place through vertical transmission (p_N(Vert), Equation 18) (Figure 1(A)).

For the default parameterisation (Figure 1(B)), the expected profits for the non-controllers (P_N) are always higher than that of the population, so they should always have a zero probability of switching strategy (Figure 1(C)). Conversely, P_C is always below the population average, and thus they should always have a non-zero probability of switching. This results in an equilibrium where there are no controllers present, as all of them have abandoned using the unprofitable CSS. However, by reducing the cost of clean seed ($\phi = 0.125$, equivalent to a 50% subsidy), at around $I = 411$ fields control becomes more profitable and non-controllers have a non-zero probability of switching strategy (Figure 1(E)). As at this point controllers cannot switch strategy, there is an "all control" equilibrium. The point at which $P_N = P_C$ (i.e. when $I \approx 411$, which is the total number of infected fields in the system and does not describe the number of infected controllers and non-controllers) leads to an equilibrium where both control strategies are present.

The value of η (the responsiveness of growers has no effect on the equilibrium attained by the

growers in the "strategy vs. population" models, as the equilibrium is based only on the relative values of P_C, P_N and P, though it did affect the time it took to reach equilibrium (Figure 1(D)). This also holds true for the "strategy vs. alternative" model.

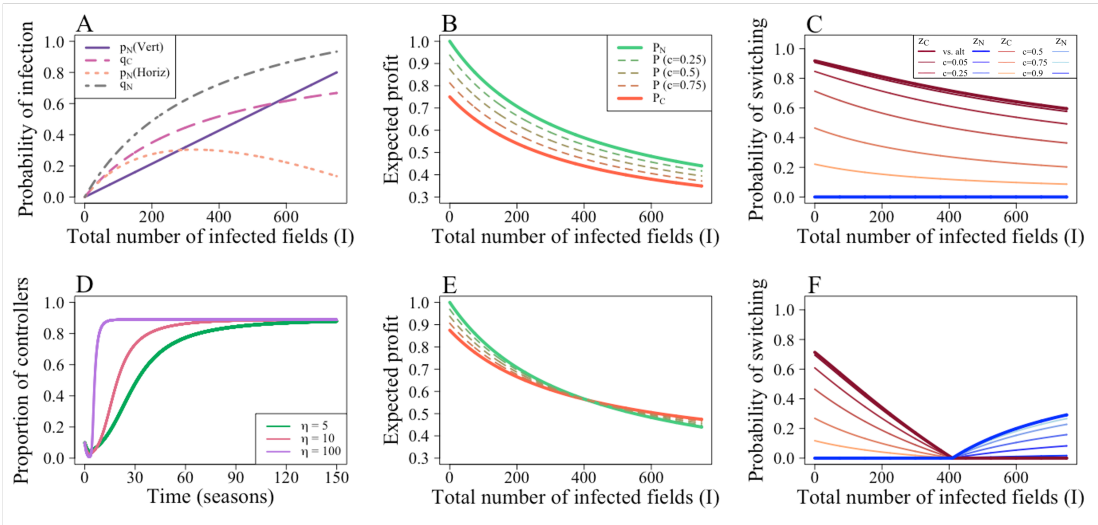

Figure 1: Probability of infection, responsiveness of growers, expected profits, and switching terms for the "strategy vs. population" model. (A) For the default parameterisation, the probability of infection increases as the number of infected growers increase (q_N and q_C), though the probability non-controllers will be infected via horizontal transmission ($p_{N(\text{Horiz})}$) falls as I increases. (B) The expected profits for controllers, non-controllers and the population for different proportions of controllers in the population (c). For the default parameters, $P_C < P$ for all values of I. (C) Switching probabilities for controllers and non-controllers for different c. As $P_C < P$, controllers should always have a non-zero probability of switching strategy, whereas non-controllers should never start to control. (D) The responsiveness of growers (η) does not affect the final equilibrium values, though it does impact the time it takes to reach equilibrium. (E) shows the expected profits for controllers, non-controllers and the population when the cost of control $\phi, = 0.125$. Now, at high levels of I, it is profitable to control. Where $P_C = P_N$, both strategies will be present at equilibrium. (F) For $\phi = 0.125$, for low values of I controllers should switch strategy, but as I increases non-controllers should have a non-zero probability of switching into the clean seed system. For this parameter set, the equilibrium value of I is therefore 411, with 349 infected controllers and 63 infected non-controllers.

B Appendix 2: Spatial-stochastic Model

As a supplementary question, we investigated whether the results we obtained with the "grower vs. alternative" assessment of profitability were robust to the addition of space and stochasticity.

B.1 Methods

B.1.1 Host landscape and initial conditions

Disease spreads on a 1,000 Ha square landscape containing $N = 750$ identically-sized fields (a similar density to [13]), placed at uniform random locations. All simulations are started with 75 growers using clean seed (i.e. 10% of the total population), distributed randomly across the landscape. A random selection of eight non-controllers is set to be infected initially ($\approx 1\%$ total population). The location of the growers was kept constant between simulation runs, though the growers that were either initially infected or used the clean seed system differed.

B.1.2 Dispersal and force of infection

Since horizontal transmission is a consequence of the movement of whitefly, the probability that a particular infected field leads to infection of a susceptible field depends upon the distance d between them. We model this using an exponential dispersal kernel

$$K(d, \alpha) = \frac{1}{2\pi\alpha^2} \exp\left(-d/\alpha\right), \tag{1}$$

in which α is the dispersal scale [69, 70]. This sets the force of infection, Γ_k, upon the field belonging to grower k

$$\Gamma_k = \beta_S \sum_{\substack{l \neq k \\ l \text{ is infected}}} K(d_{k,l}, \alpha), \tag{2}$$

in which the sum runs over all infected fields (l), $d_{k,l}$ is the distance between fields k and l, and β_S is the effective rate of infection in the spatial version of our model (Appendix 2 Table 4). The force of infection sets the rate at which grower k becomes infected; it varies in both time and space (since it depends not only on the instantaneous number of infected fields, but also on their spatial arrangement).

Table 4: Summary of parameter values and initial conditions required for the spatial-stochastic model.

Parameter	Meaning	Value	Reference
$1/\gamma$	Length of the growing season	300 days	[13]; [46]
η	Responsiveness of growers	10	Assumed (see main text)
Y	Maximum yield	1	All values scaled relative to yield
L	Loss due to infection	0.6	[38]; [47]
ϕ	Cost of control	0.25	[13]
N	Total number of fields/growers	750	Illustrative
α	Dispersal scale parameter for whitefly	150 m	[13]; [71]
β_S	Effective rate of secondary infection (spatial)	110 day^{-1} field^{-1}	Calibrated to [13]
$S_C(0)$	Initial proportion of susceptible controllers	$0.1N$	Illustrative
$I_C(0)$	Initial proportion of infected controllers	0	Illustrative
$S_N(0)$	Initial proportion of susceptible non-controllers	$0.88.93N$	Illustrative
$I_N(0)$	Initial proportion of infected non-controllers	$0.0107N$	Illustrative

B.1.3 Model dynamics

Each individual grower is characterised by their current strategy as adopted on planting (i.e. controller or non-controller), as well as their current infection status (i.e. susceptible or infected). Between harvesting and replanting events, dynamics are independent of the strategy adopted. We simulate the model using Gillespie's algorithm [72].

The only event that can affect an infected grower (i.e. class I_C or I_N) is harvesting; this occurs at rate γ. As well as harvesting, susceptible growers (i.e class S_C or S_N) can also become horizontally infected. This occurs, for grower k, at rate Γ_k (Equation 2). As noted above, the force of infection not only depends on time, but also takes different values for each grower.

B.1.4 Replanting and grower behaviour

Replanting is assumed to occur immediately after harvesting, and potentially leads to changes in both strategy and infection status. Different growers make different assessments of the probability

of horizontal infection next season depending on their instantaneous force of infection, with

$$\begin{aligned}\text{Estimated probability of} \\ \text{horizontal infection for grower } k \end{aligned} &= \frac{\text{Force of infection on grower } k}{\text{Force of infection on grower } k + \text{Harvesting rate}}, \quad (3)$$

$$= \frac{\Gamma_k}{\Gamma_k + \gamma}, \quad (4)$$

We assume growers make decisions according to spatial analogues of Equations 17 and 20 in the main text:

$$q_C^k = \text{Grower } k\text{'s estimate of the probability of infection next season if control is adopted}, \quad (5)$$

$$= \frac{\Gamma_k}{\Gamma_k + \gamma}, \quad (6)$$

and

$$q_N^k = \text{Grower } k\text{'s estimate of the probability of infection next season if control is not adopted},$$

$$= \frac{p(I_C + I_N)}{N} + \left(\frac{\Gamma_k}{\Gamma_k + \gamma}\right)\left(1 - \frac{p(I_C + I_N)}{N}\right). \quad (7)$$

All other aspects of the behavioural model follow essentially unchanged, although the switching terms now enter the model via a single Bernoulli trial ([73]) at the time of planting to determine whether a grower switches strategy. Whether or not non-controllers become vertically infected is also simulated using a single Bernoulli trial at the time of planting (rather than via a systematic alteration to a rate parameter as in Equations 5-8 in the main text).

B.1.5 Misestimating parameters

We also investigated misestimation of epidemiological parameters in the spatial model, focusing on the dispersal scale (α) and and the rate of horizontal transmission, β_S:

$$q_\alpha = \nu_\alpha \alpha, \tag{8}$$

$$q_{\beta_S} = \nu_{\beta_S} \beta_S. \tag{9}$$

The exponential dispersal kernel ($\tilde{K}(d, q_\alpha)$) used in growers' decision making is now given by:

$$\tilde{K}(d, q_\alpha) = \frac{1}{2\pi q_\alpha^2} \exp\left(-d/q_\alpha\right), \tag{10}$$

which sets the estimated force of infection, $\tilde{\Gamma}_k$, upon the field belonging to grower k via

$$\tilde{\Gamma}_k = q_{\beta_S} \sum_{\substack{l \neq k \\ l \text{ is infected}}} \tilde{K}(d_{k,l}, q_\alpha). \tag{11}$$

B.1.6 Parameterisation

The majority of parameters in the spatial model took the same values to those in the simpler non-spatial model (Table 1) that was introduced in the main text. However, following [13] and based on dispersal of *B. tabaci* reported by [74] and [71], we set the scale parameter of our dispersal kernel to be $\alpha = 150$ m. The move to a spatially-explicit model required a re-scaling of the rate of secondary infection, which we set to $\beta = 110$ day^{-1} field^{-1}. This value was again chosen to ensure that, after 10 seasons, on average 50% of fields would be infected.

B.2 Results

Under the default parameterisation, no grower controls (Figure 1(A)), though providing a 50% subsidy encouraged participation in the CSS (Appendix 2 Figure 1(B)). Disease spreads rapidly between fields, both due to the high values of β and ϕ (which discourage control) and the non-spatial aspect of trade, which allows disease to spread far away from its initial source.

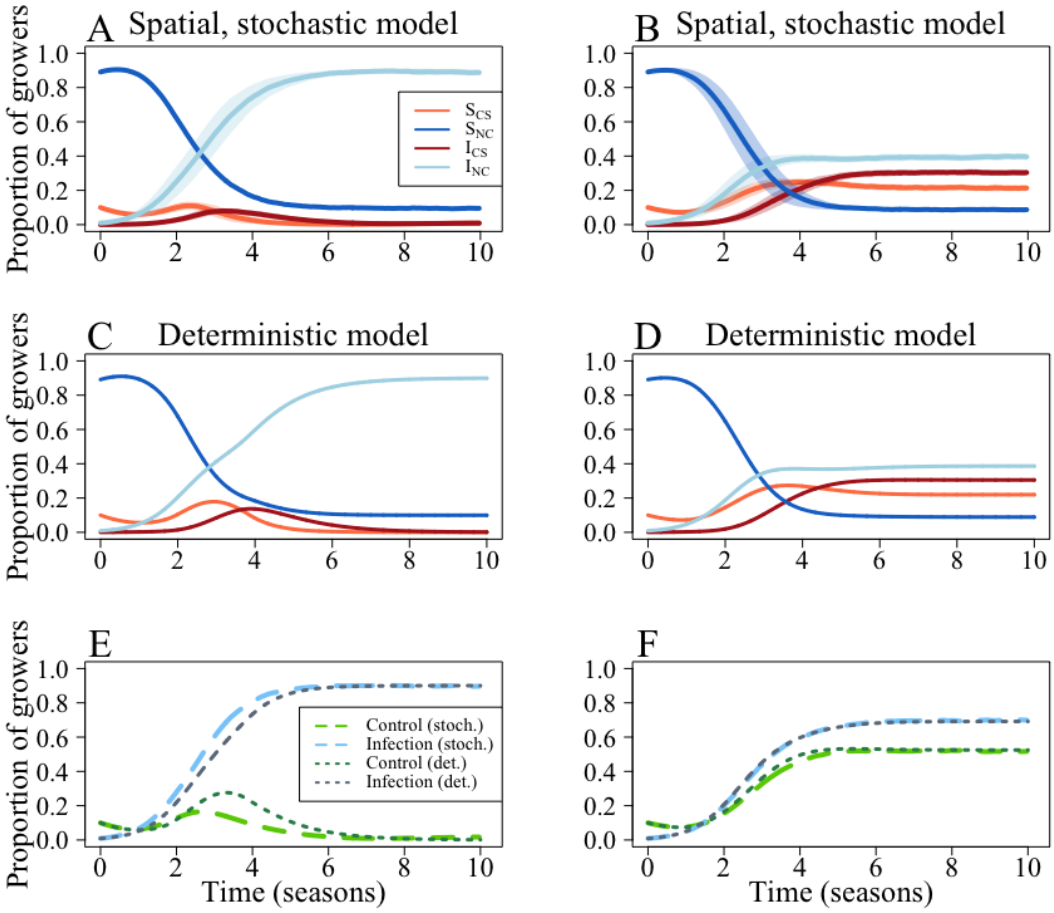

Figure 1: Comparison of default behaviour of the "grower vs. alternative" spatial-stochastic and deterministic models. (A) Dynamics for spatial model. As with the deterministic model (C), under the default parameterisation no growers use the CSS after 10 seasons. Adding a subsidy in (B) and (D) allows for the two-strategy equilibrium. The figures show the mean for 100 runs of each model, and the error bars show one standard deviation. The equilibrium values and dynamics for spatial ((A) and (C)) and non-spatial models ((B) and (D)) are very similar in both cases, emphasised in (E) and (F), which show the proportions controlling and infected for the spatial-stochastic ("stoch.") and deterministic ("det.") models.

The results of our spatial model are comparable with the deterministic version of the model (Appendix 2 Figure 2). For equivalent rates of horizontal transmission, the equilibrium values for the deterministic model are similar to the values attained after 50 seasons in the spatial-stochastic model (Appendix 2 Figure 2 (A) and (B)). Disease extinction in the stochastic model was rare; it occurred in every simulation when $\beta_S = 0$ day^{-1} field^{-1}, and around 40% of the time when $\beta_S = 40$ day^{-1} field^{-1}. Above $\beta_S = 40$ day^{-1} field^{-1}, disease extinction was uncommon.

As in Figure 6 of the main text, which shows the analogous result for the non-spatial deterministic model, the kinks in these graphs are a result of susceptible controllers (S_C) growers ceasing to switch strategy (around $\beta_s = 50$ day^{-1} field^{-1} and $\beta = 0.003$ day^{-1} in Appendix 2 Figure 2(A) and (B)). As the infection pressure is increasing, these susceptible growers then become infected and switch into the non-control strategy, causing an overall decrease in controllers.

In both cases, when ϕ is varied, at around $\phi = 0.19$, growers should stop using the control strategy (Figure 2(C) and (D)).

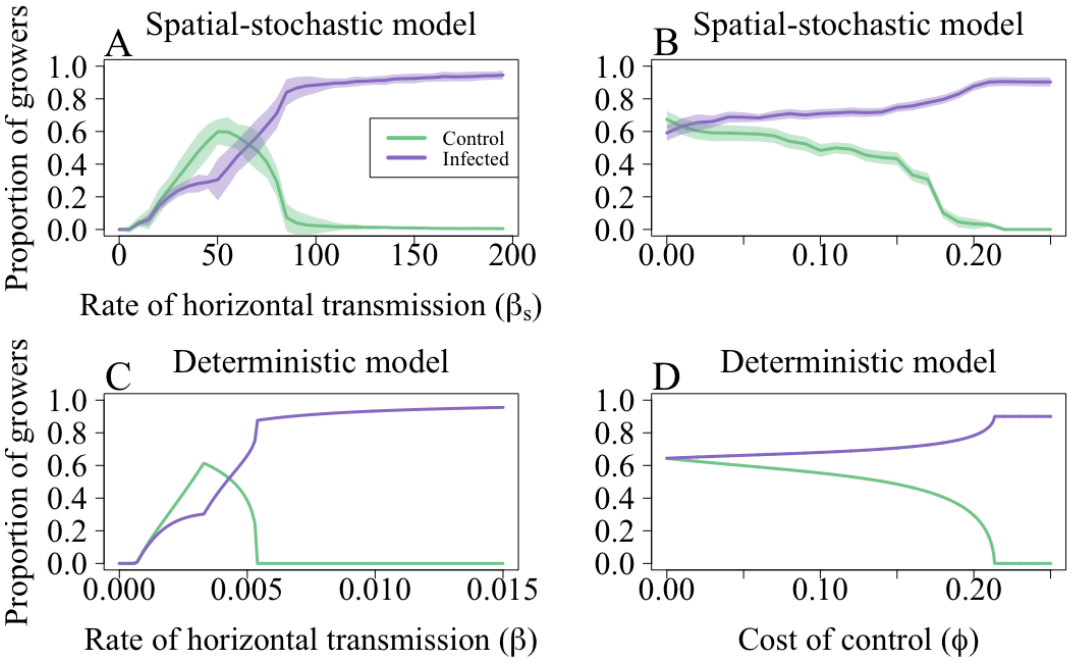

Figure 2: Response to changes in the rate of horizontal transmission and cost of control for the "grower vs. alternative" models. (A) and (C) The proportion of controllers and infected fields after 50 seasons for the spatial-stochastic model and (B) and (D) the equilibrium values of control ($S_C + I_C$) and infection ($I_N + I_C$) for the deterministic model. Aside from the parameters being scanned over, the default parameters are used (Table 1 and 4). The results in (A) and (C) closely align with the equilibrium values in the non-spatial model ((B) and (D)), indicating that our results are robust to spatial and stochastic effects. In (A) and (C) the means over 100 runs and the error bars show one standard deviation around the mean.

B.2.1 Effect of systematic uncertainty

Growers misestimating the dispersal scale for the whitefly vector (α) has a similar effect to when growers misestimated β in the deterministic model. As the perceived value of α (q_α, given by $q_\alpha = \nu_\alpha \alpha$, Table 4) increases, fewer growers participate in the CSS as they estimated that they would be paying the dual penalty of the cost of control and loss due to disease (Figure 3).

When growers misestimate the rate of horizontal transmission, β_S, the pattern of CSS use again resembles that of the deterministic model (Figure 6(A)). At low values of the perceived value of β (q_{β_S}, with $q_{\beta_S} = \nu_{\beta_S}\beta_S$, Table 4), more growers use the control scheme as they believe that they are unlikely to be infected. As q_{β_S} increases, growers believe that they are likely to pay the dual penalty of the cost of control, ϕ, and the loss due to infection, L, and therefore abandon the CSS.

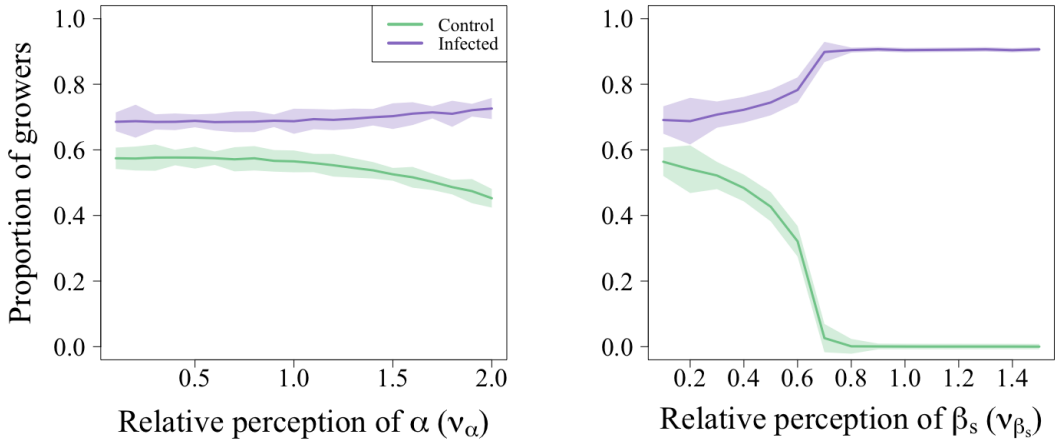

Figure 3: Effect of systematic misestimation in the spatial-stochastic model. (A) For the default parameters, as the perceptions of the dispersal scale for the whitefly vector (ν_α) increase, fewer growers use the control scheme as they estimate that they would likely end up infected. (B) As perceptions of the rate of horizontal transmission increase (ν_β), fewer growers use the CSS in a pattern that matches that seen in Figure 6(A) in the main text. The mean values were calculated over 100 runs are the error bars show one standard deviation around the mean.

B.2.2 Spatial spread of disease.

Appendix 2 Figure 4 shows the spatial component of disease spread in the spatial-stochastic model, restricting attention to the case when there is only horizontal transmission ($p = 0$). Disease then

spreads in a much more "wavelike" pattern, as expected from the thin-tailed exponential dispersal kernel as adopted for whitefly. As there was no vertical transmission, we also removed "control" from the strategy set of the growers, as there was no benefit to controlling for disease and thus the strategy would quickly disappear from the population. Unlike the other models, we also start our epidemic with a cluster of infected fields. Making these changes allows us to focus on the underlying dynamics of the model.

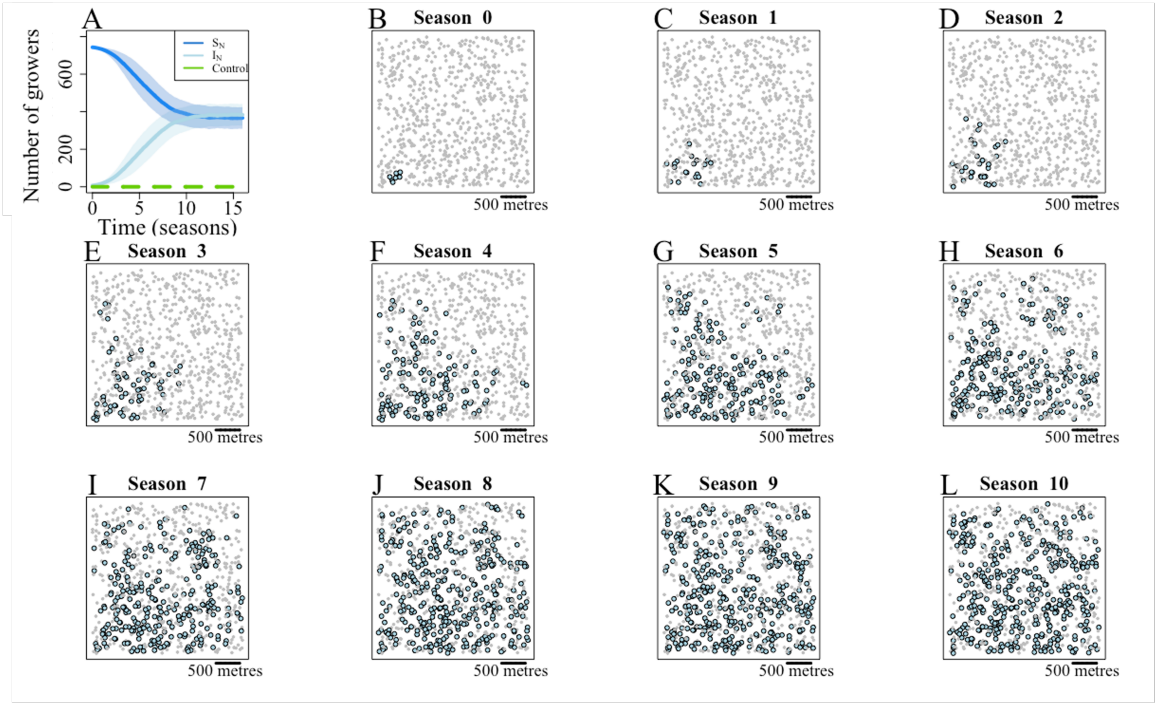

Figure 4: Spatial spread of infection in spatial-stochastic model. To emphasise the spatial component of infection, vertical transmission has been removed from the model (i.e. $p = 0$). Additionally, we have removed the "control" strategy from the growers as, without vertical transmission, this would have disappeared from the population within 5 seasons. The light blue dots show infected fields, whilst the grey dots are uninfected fields.

B.3 Discussion

The results for this model were comparable with those for the equivalent deterministic model (Appendix 2 Figure 2), demonstrating the robustness of our results to stochastic and spatial effects. This was in part due to the non-spatial aspect of trade, as it allowed for transmission across the relatively small landscape considered here [42]. When trade was not included (i.e. $p = 0$) the disease

spread across the landscape in a more "wavelike" pattern as expected from the thin-tailed dispersal kernel for the whitefly ([75]; Appendix 2 Figure 4). However, when trade occurs via a market or central organisation the majority of transactions occur over a scale larger than our landscape (a square with sides 3.16 km), with around 70 % of transactions occur over a scale of 10-50 km [14]. For more informal trade settings, with growers interacting with each other, a proximity-based kernel may be more appropriate ([42], [13], though both of these were modelled over a larger landscape than ours). However, such exchanges are hard to parameterise as the probability of exchanging planting material is highly variable across settings (varying between 30 - 92.51% over unspecified spatial scales; [76], [77], [78], [79] and [80]).

C Appendix 3: Mathematical derivations

C.1 The basic reproductive number (R_0) for the "strategy vs." models.

To find the value for R_0, we calculated the next-generation matrix (NGM; [81]). This relies on the decomposing a linearised version of the model into two matrices: the first contains the terms relating to disease transmission (matrix J_F) whilst the second has terms relating to non-epidemiological transitions between states (matrix J_V). The NGM, K, is given by $K = J_F J_V^{-1}$ [82]. To shorten the notation in what follows, we introduce the following function of state variables:

$$\Gamma = S_C z_{SC} + I_C z_{IC} + S_N(1 - z_{SN}) + I_N(1 - z_{IN}). \tag{1}$$

We focus only on the infected compartments in the general model, leading to:

$$\frac{dI_C}{dt} = \beta S_C(I_C + I_N) - \gamma I_C, \tag{2}$$

$$\frac{dI_N}{dt} = \gamma\Gamma\left(\frac{p(I_C + I_N)}{N}\right) + \beta S_N(I_C + I_N) - \gamma I_N. \tag{3}$$

The disease-free equilibrium (DFE) is given by $(S_C, I_C, S_N, I_N) = (0, 0, N, 0)$. Given that we have restricted the model to just the infected compartments, the DFE can also be written as $(I_C, I_N) = (0, 0)$

The matrix of rates at which new infections occur is

$$F = \begin{bmatrix} \beta S_C(I_C + I_N) \\ \gamma\Gamma\left(\frac{p(I_C+I_N)}{N}\right) + \beta S_N(I_C + I_N) \end{bmatrix}. \tag{4}$$

The matrix of rates at which infections are removed is

$$V = \begin{bmatrix} \gamma I_C \\ \gamma I_N \end{bmatrix}. \tag{5}$$

The Jacobians for these matrices are:

$$J_F = \begin{bmatrix} \beta S_C & \beta S_C \\ \gamma\Gamma\left(\frac{p}{N}\right) + \gamma\frac{\partial\Gamma}{\partial I_C}\left(\frac{p(I_C+I_N)}{N}\right) + \beta S_N & \gamma\Gamma\left(\frac{p}{N}\right) + \gamma\frac{\partial\Gamma}{\partial I_N}\left(\frac{p(I_C+I_N)}{N}\right) + \beta S_N \end{bmatrix}, \tag{6}$$

and

$$J_V = \begin{bmatrix} \gamma & 0 \\ 0 & \gamma \end{bmatrix}. \tag{7}$$

We need to evaluate the Jacobians at the DFE. Note that $\gamma\Gamma$ is the net rate at which fields in which there is no control are planted, and so at the DFE, $\Gamma = N$. Note too that, since both are multiplied by $\frac{p(I_C+I_N)}{N}$ (which is zero at the DFE), the partial derivatives and $\frac{\partial\Gamma}{\partial I_C}$ and $\frac{\partial\Gamma}{\partial I_N}$ do not need to be calculated.

At the DFE, the J_F becomes

$$J_F = \begin{bmatrix} 0 & 0 \\ \gamma p + \beta N & \gamma p + \beta N \end{bmatrix}, \tag{8}$$

and J_V^{-1} is given by:

$$J_V^{-1} = \begin{bmatrix} 1/\gamma & 0 \\ 0 & 1/\gamma \end{bmatrix}. \tag{9}$$

The NGM, $K = J_F J_V^{-1}$, is therefore given by:

$$\begin{bmatrix} 0 & 0 \\ p + \frac{\beta N}{\gamma} & p + \frac{\beta N}{\gamma} \end{bmatrix} \tag{10}$$

The dominant eigenvalue for this matrix is $\lambda_1 = \frac{\beta N}{\gamma} + p$, which gives the R_0 for the system. This can be further broken down into distinct components corresponding to horizontal ($R_0^H = \frac{\beta N}{\gamma}$) and vertical ($R_0^V = p$) transmission [52].

C.2 Stability of equilibria in the "strategy vs. population" model.

There are four equilibria, which can be distinguished by the presence of disease and the proportion of growers controlling at equilibrium.

- **Disease-free equilibrium** at which the disease is not able to spread even when there is no control via clean seed

$$(S_C, I_C, S_N, I_N) = (0, 0, N, 0). \tag{11}$$

- **Control-free, disease-endemic equilibrium** at which the disease is able to spread in the

absence of control, but nevertheless no grower controls

$$(S_C, I_C, S_N, I_N) = \left(0, 0, \frac{\gamma(1-p)}{\beta}, \frac{\beta N - \gamma(1-p)}{\beta}\right). \tag{12}$$

- **All-control, disease-endemic equilibrium** at which all growers control, but nevertheless disease is still present in the system

$$(S_C, I_C, S_N, I_N) = \left(\frac{\gamma}{\beta}, \frac{\beta N - \gamma}{\beta}, 0, 0\right) \tag{13}$$

- **Two-strategy, disease-endemic equilibrium** at which both disease and control equilibrate at some intermediate level, with

$$S_C = \frac{N(R_0(\gamma pL - \beta\phi N) - \gamma pL)}{\gamma p^2 L}, \tag{14}$$

$$I_C = \frac{\beta\phi N^2(R_0(\beta\phi N - \gamma pL) + \gamma pL)}{\gamma p^2 L(\beta\phi N - \gamma pL)}, \tag{15}$$

$$S_N = \frac{N(\phi R_0 - pL)(\beta N(\beta\phi N - \gamma pL) + \gamma^2 pL)}{\gamma p^2 L(\beta\phi N - \gamma pL)}, \tag{16}$$

$$I_N = \frac{-\phi N R_0(\beta N(\beta\phi N - \gamma Lp) + \gamma^2 Lp)}{\gamma p^2 L(\beta\phi N - \gamma Lp)}. \tag{17}$$

We also note that none of Equations 11, 12, 12 or 14 - 17 have a dependence on the responsiveness of growers, η. Thus, the equilibrium attained is independent of this parameter, though it does affect the dynamics approaching equilibrium (Appendix 2 Figure 1(D)).

The stability of each of these equilibria will be discussed in turn.

C.2.1 Disease-free equilibrium.

We determined the conditions for stability of each equilibrium in the "strategy vs." models by first evaluating the Jacobian matrix for the system at each possible equilibrium and then determining

the eigenvalues for the matrix. The system can be reduced to three state variables, as $N = S_C + I_C + S_N + I_N$. The therefore becomes:

$$\frac{dS_C}{dt} = \gamma\left(S_C(1-z_{SC}) + I_C(1-z_{IC}) + (N - S_C - I_C)z_{IN}\right) - \beta S_C\left(I_C + I_N\right) - \gamma S_C, \tag{18}$$

$$\frac{dI_C}{dt} = \beta S_C\left(I_C + I_N\right) - \gamma I_C, \tag{19}$$

$$\frac{dI_N}{dt} = \gamma\left(S_C z_{SC} + I_C z_{IC} + (N - S_C - I_C)(1 - z_{IN})\right)\left(\frac{p(I_C + I_N)}{N}\right) + \beta(N - S_C - I_C - I_N)\left(I_C + I_N\right) - \gamma I_N. \tag{20}$$

where the switching terms are given by:

$$z_{SC} = z_{IC} = \max\left(0, 1 - e^{-\eta(P - P_C)}\right), \tag{21}$$

$$z_{SN} = z_{IN} = \max\left(0, 1 - e^{-\eta(P - P_N)}\right). \tag{22}$$

The DFE is given by $(S_C, I_C, I_N) = (0, 0, 0)$. The Jacobian matrix evaluated with these values is given by:

$$\begin{bmatrix} -\gamma(1 - \exp(-\eta\phi) + \eta\phi) & \gamma(\exp(-\eta\phi) - \eta\phi) & 0 \\ 0 & -\gamma & 0 \\ 0 & \beta N + \gamma p & \beta N + \gamma p - \gamma \end{bmatrix}$$

From this, we can see the first eigenvalue is $\beta N - \gamma + \gamma p$. The 2×2 matrix that remains is given by:

$$\begin{bmatrix} -\gamma(1 - \exp(-\eta\phi) + \eta\phi) & \gamma(\exp(-\eta\phi) - \eta\phi) \\ 0 & -\gamma \end{bmatrix}$$

This matrix is upper triangular, so the eigenvalues are given by the diagonal elements. We can see that $-\gamma$ and $-\gamma(1 - \exp(-\eta\phi) + \phi)$ are the remaining eigenvalues for the matrix.

As $-\gamma$ and $-\gamma(1-\exp(-\eta\phi)+\phi)$ are always negative, the stability of this equilibrium is dependent upon $\beta N - \gamma + \gamma p < 0$. This corresponds to the R_0 found using the NGM.

C.2.2 Disease-endemic, all control equilibrium.

The disease-endemic, all control equilibrium is given by: $(S_C, I_C, I_N) = \left(\frac{\gamma}{\beta}, \frac{\beta N - \gamma}{\beta}, 0\right)$. As there are no non-controllers at this equilibrium, $P_C = P$ (Equation 23 in the main text).

The Jacobian matrix evaluated using these values is given by:

$$\begin{bmatrix} \frac{\partial \dot{S_C}}{\partial S_C} & \frac{\partial \dot{S_C}}{\partial I_C} & -\gamma \\ \beta N - \gamma & 0 & \gamma \\ \frac{\partial \dot{I_N}}{\partial S_C} & \frac{\partial \dot{I_N}}{\partial I_C} & -\gamma \end{bmatrix} \tag{23}$$

with:

$$\frac{\partial \dot{S_C}}{\partial S_C} = \gamma \left(\exp\left(\frac{-\eta(\beta N(\gamma p L - \beta \phi N) - \gamma^2 p L)}{\beta^2 N^2} \right) \right) - \frac{-\eta(\beta N(\gamma p L - \beta \phi N) - \gamma^2 p L)}{\beta^2 N^2} - \beta N \tag{24}$$

$$\frac{\partial \dot{S_C}}{\partial I_C} = -\gamma \left(1 - \exp\left(\frac{-\eta(\beta N(\gamma p L - \beta \phi N) - \gamma^2 p L)}{\beta^2 N^2} \right) \right) - \frac{-\eta(\beta N(\gamma p L - \beta \phi N) - \gamma^2 p L)}{\beta^2 N^2} \tag{25}$$

$$\frac{\partial \dot{I_N}}{\partial S_C} = \beta N + \gamma - \frac{\gamma p(\beta N - \gamma)}{\beta N} \left(\exp\left(\frac{-\eta(\beta N(\gamma p L - \beta \phi N) - \gamma^2 p L)}{\beta^2 N^2} \right) - \frac{-\eta(\beta N(\gamma p L - \beta \phi N) - \gamma^2 p L)}{\beta^2 N^2} \right) \tag{26}$$

$$\frac{\partial \dot{I_N}}{\partial I_C} = \frac{\partial \dot{I_N}}{\partial S_C}. \tag{27}$$

We can see that

$$\frac{\partial \dot{S_C}}{\partial S_C} = \frac{\partial \dot{S_C}}{\partial I_C} - \gamma + \beta N. \tag{28}$$

The characteristic equation for Matrix 23 is given by

$$-(\lambda + \gamma) \left(\lambda^2 - \lambda \left(\frac{\partial \dot{S_C}}{\partial I_C} - \gamma + \beta N \right) - (\beta N - \gamma) \left(\frac{\partial \dot{S_C}}{\partial I_C} \right) \right) \tag{29}$$

Solving this equation, we find the eigenvalues to be $-\gamma$, $\beta N - \gamma$ and

$$\frac{\partial \dot{S_C}}{\partial I_C} = -\gamma \left(1 - \exp\left(\frac{-\eta(N\beta(\gamma pL - \beta\phi N) - \gamma^2 pL)}{\beta^2 N^2}\right) + \frac{\eta(\beta N(\gamma pL - \beta\phi N) - \gamma^2 pL)}{\beta^2 N^2}\right). \tag{30}$$

As $-\gamma$ is always negative, the stability depends on the remaining eigenvalues. The second eigenvalue is always negative for $\frac{\beta N}{\gamma} > 1$, which is also the R_0 for horizontal transmission (R_0^H). Given that this expression can be re-written as:

$$\frac{\partial \dot{S_C}}{\partial I_C} = -\gamma \left(1 - \exp\left(-x\right) + x\right), \tag{31}$$

where $x = \frac{\eta(\beta N(\gamma pL - \beta\phi N) - \gamma^2 pL)}{\beta^2 N^2}$, the final eigenvalue is negative once when the term in the exponent (x) is negative:

$$\frac{-\eta(\beta N(\beta\phi N - \gamma pL) + \gamma^2 pL)}{\beta N} < 0 \tag{32}$$

$$(\beta N(\beta\phi N - \gamma pL) + \gamma^2 pL) < 0 \tag{33}$$

$$\frac{\beta N(\beta\phi N - \gamma pL)}{\gamma} < \gamma pL \tag{34}$$

$$\gamma pL \left(\frac{\gamma - \beta N}{\beta^2 N^2}\right) > \phi \tag{35}$$

$$pL \left(\frac{\gamma - \beta N}{\beta N}\right) \left(\frac{\gamma}{\beta N}\right) > \phi. \tag{36}$$

For clarity, this can be re-arranged as follows:

$$\begin{array}{c}\text{Cost of}\\ \text{control}\end{array} < \begin{array}{c}\text{Loss due to}\\ \text{disease}\end{array} \times \begin{array}{c}\text{Probability of}\\ \text{vertical transmission}\end{array} \times \begin{array}{c}\text{Proportion of}\\ \text{infected fields}\end{array} \times \left(1 - \begin{array}{c}\text{Probability of}\\ \text{horizontal transmission}\end{array}\right),$$

$$\phi < Lp \left(\frac{\beta N - \gamma}{\beta N}\right) \left(1 - \frac{\beta N - \gamma}{\beta N}\right). \tag{37}$$

That is, for the "all control" equilibrium to be stable, $R_0^H > 1$ and the cost of control (ϕ) must be less than the expected losses of infected non-controllers (L).

C.2.3 Disease-endemic, control-free equilibrium.

The disease-endemic, control-free equilibrium is given by: $(S_C, I_C, I_N) = \left(0, 0, \frac{\beta N - \gamma(1-p)}{\beta}\right)$. Additionally, as there are no controllers at this equilibrium, $P_N = P$ (Equation 23 in the main text). The entries for the Jacobian matrix evaluated at the disease-endemic, no control equilibrium are given as follows:

$$\begin{bmatrix} \frac{\partial \dot{S_C}}{\partial S_C} & \frac{\partial \dot{S_C}}{\partial I_C} & 0 \\ \gamma(R_0 - 1) & -\gamma & 0 \\ \frac{\partial \dot{I_N}}{\partial S_C}^* & \frac{\partial \dot{I_N}}{\partial I_C}^* & \gamma(R_0 - 1) \end{bmatrix} \tag{38}$$

Entries marked with "*" are not written out in full as they are not needed for further analysis. Remaining entries are given as follows:

$$\frac{\partial \dot{S_C}}{\partial S_C} = \gamma \left(\exp\left(-\frac{\eta(\gamma R_0(\beta\phi N - \gamma p L) + \gamma^2 p L)}{\beta\gamma N R_0} \right) - \frac{\gamma\eta(R_0(\beta\phi N - \gamma p L) + \gamma p L)}{\beta N R_0} - R_0 \right) \tag{39}$$

$$\frac{\partial \dot{S_C}}{\partial I_C} = \gamma \left(\exp\left(-\frac{\eta(\gamma R_0(\beta\phi N - \gamma p L) + \gamma^2 p L)}{\beta\gamma N R_0} \right) - \frac{\gamma\eta(R_0(\beta\phi N - \gamma p L) + \gamma p L)}{\beta N R_0} \right) \tag{40}$$

from which we can see that

$$\frac{\partial \dot{S_C}}{\partial S_C} = \frac{\partial \dot{S_C}}{\partial I_C} - \gamma R_0. \tag{41}$$

From this matrix, we can see that the first eigenvalue is given by $\gamma(R_0 - 1)$. This is negative for $R_0 > 1$.

The remaining eigenvalues can be found by solving the equation:

$$\lambda^2 - a_1\lambda + a_2 = 0 \tag{42}$$

where

$$a_1 = \frac{\partial \dot{S_C}}{\partial I_C} - \gamma(R_0 - 1) \tag{43}$$

$$a_2 = -\gamma R_0 \left(\frac{\partial \dot{S_C}}{\partial I_C} - \gamma \right). \tag{44}$$

Solving this, we find the remaining eigenvalues to be: $-\gamma R_0$, and

$$\gamma \left(\exp\left(-\frac{\eta(\gamma R_0(\beta\phi N - \gamma p L) + \gamma^2 p L)}{\beta\gamma N R_0} \right) - 1 + \frac{-\eta(R_0((\beta\phi N - \gamma p L)) - \gamma\eta p L)}{\beta N} \right). \tag{45}$$

Clearly, $-\gamma R_0$ is always negative. The final eigenvalue is negative once

$$1 - \exp\left(-\frac{\eta(\gamma R_0(\beta\phi N - \gamma p L) + \gamma^2 p L)}{\beta\gamma N R_0} \right) < \frac{-\eta(R_0((\beta\phi N - \gamma p L)) - \gamma\eta p L)}{\beta N} \tag{46}$$

That is, the eigenvalue is negative once the term in the exponent is negative. This can be rearranged as follows:

$$Lp \left(\frac{\beta N - \gamma(1-p)}{\beta N} \right) \left(\frac{\gamma}{\beta N + \gamma p} \right) < \phi \tag{47}$$

For clarity, this can be re-written in terms of the probability of infection:

$$\begin{array}{c}\text{Loss due to}\\ \text{disease}\end{array} \times \begin{array}{c}\text{Probability of}\\ \text{vertical transmission}\end{array} \times \begin{array}{c}\text{Proportion of}\\ \text{infected fields}\end{array} \times \left(1 - \begin{array}{c}\text{Probability of}\\ \text{horizontal transmission.}\end{array} \right) < \begin{array}{c}\text{Cost of}\\ \text{control,}\end{array} \tag{48}$$

$$Lp \left(\frac{\beta N - \gamma(1-p)}{\beta N} \right) \left(1 - \frac{\beta N - \gamma(1-p)}{\beta N + \gamma p} \right) < \phi. \tag{49}$$

Therefore, the control-free, disease-endemic equilibrium is stable for $R_0 > 1$ and when the expected losses due to disease for an infected non-controller are less than the cost of control.

C.3 Stability of two-strategy equilibrium in the "strategy. vs" models.

Our "strategy vs." models are discontinuous, as the switching terms take different forms depending on the relative values of the expected profits for each strategy and the population. Due to this discontinuity, the stability of the two-strategy equilibrium for the "strategy vs. population" model cannot be assessed by linearising around the equilibrium. We instead used a numerical approach to assess the stability of this equilibrium. We first generated a parameter set by sampling parameters from a plausible range (given in Appendix 3 Table 5). We then ran the model for each parameter set and found the equilibrium values for each state variable ($E^* = (S_U^*, I_U^*, S_C^*, I_C^*)$, taken after 500 seasons). Then, for the same parameter set, the model was run with a set of 10,000 different initial conditions again for 500 seasons. The difference between the final equilibrium value for each state variable and E^* was calculated. If the difference was greater than 10^{-8}, it was deemed large enough to conclude that the equilibrium was unstable based on different initial conditions. We repeated this for 5,000 different parameter sets and, in each instance, the equilibrium attained was that expected for the parameter set.

Table 5: Range of values used for parameters when evaluating the stability of the "grower vs." models.

Parameter	Meaning	Range of values
$1/\gamma$	Length of the growing season	$1 - 600$ days
β	Rate of secondary infection	$0 - 2 \times 10^{-5}$ day^{-1} field^{-1}
p	Probability of getting infected cuttings	$0 - 1$
η	Responsiveness of growers	$1 - 200$
L	Loss due to infection	$0 - 1$
ϕ	Cost of control	$0 - 1$
q_j^k	Estimated value of q_j for grower k (only differs by grower in the spatial model)	Equations 6 and 7
q_α	Perceived value of the whitefly dispersal scale parameterα	Equations 8
q_{β_S}	Perceived value of horizontal transmission in the stochastic, spatial model (β_S)	Equations 9
$S_C(0)$	Initial proportion of susceptible controllers	$0 - N$
$I_C(0)$	Initial proportion of infected controllers	$0 - N$
$S_N(0)$	Initial proportion of susceptible non-controllers	$0 - N$
$I_N(0)$	Initial proportion of infected non-controllers	$0 - N$

C.4 Mutual exclusivity of equilibria

For the "no control" equilibrium to be stable, the expected losses due to vertical transmission must be less than the cost of control. That is, from Equation 49, the following must be true:

$$Lp\left(\frac{\beta N-\gamma(1-p)}{\beta N}\right)\left(1-\frac{\beta N-\gamma(1-p)}{\beta N+\gamma p}\right)<\phi \tag{50}$$

Conversely, for the "all control" equilibrium to be stable, the costs must be less than the expected losses due to vertical transmission. From Equation 37, this means that:

$$\phi<Lp\left(\frac{\beta N-\gamma}{\beta N}\right)\left(1-\frac{\beta N-\gamma}{\beta N}\right) \tag{51}$$

Thus, for both equilibria to be stable, the following condition must be met:

$$Lp\left(\frac{\beta N-\gamma(1-p)}{\beta N}\right)\left(1-\frac{\beta N-\gamma(1-p)}{\beta N+\gamma p}\right)<Lp\left(\frac{\beta N-\gamma}{\beta N}\right)\left(1-\frac{\beta N-\gamma}{\beta N}\right) \tag{52}$$

This can be simplified to:

$$\left(\frac{\beta N-\gamma(1-p)}{\beta N+\gamma p}\right)<\left(\frac{\beta N-\gamma}{\beta N}\right), \tag{53}$$

which, in turn, leaves us with the condition that $0<-\gamma^2 p$ must be true. However, as both γ and p are positive parameters, this is not possible and thus both equilibria cannot be simultaneously stable.

C.5 Numerical assessment of equilibria in the "grower vs." models.

The more complex form of these models precludes mathematical analysis, as conditions for model equilibria no longer simply depend on the difference in profits between controllers and non-controllers. For these "grower vs." models to reach a two-strategy equilibrium, the flow between the two strategies must be equal, i.e.

$$z_{IN} I_N = z_{SC} S_C + z_{IC} I_C \tag{54}$$

Consequently, we evaluated the stability of the possible equilibrium using numerical methods. Using the same method as when assessing the stability of the "two-strategy equilibrium" in the "strategy vs." models, we tested 10,000 sets of initial conditions and ran the model to equilibrium for 5,000 different parameter sets (chosen from a range of the plausible range parameters described in Appendix 3 Table 5). In no case for either of the "grower vs." models was it found that the difference between equilibrium values of the simulated model was larger than our difference threshold of 10^{-8}. We thus concluded that the final equilibrium attained was not dependent on initial conditions.

The forms of the switching terms will differ between the two "grower vs." models, leading to different equilibria for each model (Figure 3(C)-(D) in the main text). There are two important differences from the "strategy vs." models:

- The "all control" equilibrium is impossible because growers who controlled but nevertheless became infected always consider switching strategy, as they are earning the lowest possible payoff (P_{IC}, the "sucker's payoff"). As long as there is a non-zero probability of infection (which is necessary for control to be worthwhile) there will always be non-controlling growers.

- For the "grower vs. population" model, the "no control" equilibrium is only possible if all non-controllers are infected at equilibrium. For a "no control" equilibrium, the expected profit of the population must be less than or equal to P_{IN} to prevent I_N growers switching strategies (i.e. $P \leq P_{IN}$, Equation 34). At the "no control" equilibrium $P = P_N$ (Equation 23), leading

to:

$$P = \frac{P_N(S_N + I_N)}{N} \leq P_{IN} \tag{55}$$

As, in the "no control" equilibrium, $S_N + I_N = N$, this can be simplified:

$$\begin{aligned} P_N &\leq P_{IN} \\ q_N P_{IN} + (1 - q_N)P_{SN} &\leq P_{IN} \\ (1 - q_N)P_{SN} &\leq P_{IN}(1 - q_N) \\ P_{SN} &\leq P_{IN} \end{aligned} \tag{56}$$

The non-zero value of the loss due to disease (L) means that the conclusion presented in Equation 56 is impossible (see Equation 11 & 12 in the main text). Therefore, for a "no-control" equilibrium to be possible, all fields must be infected (i.e. $N = I_N$ and $q_N = 1$). In Figure 2(B) (main text), a "no control" equilibrium would require higher β and p (Figure 1).

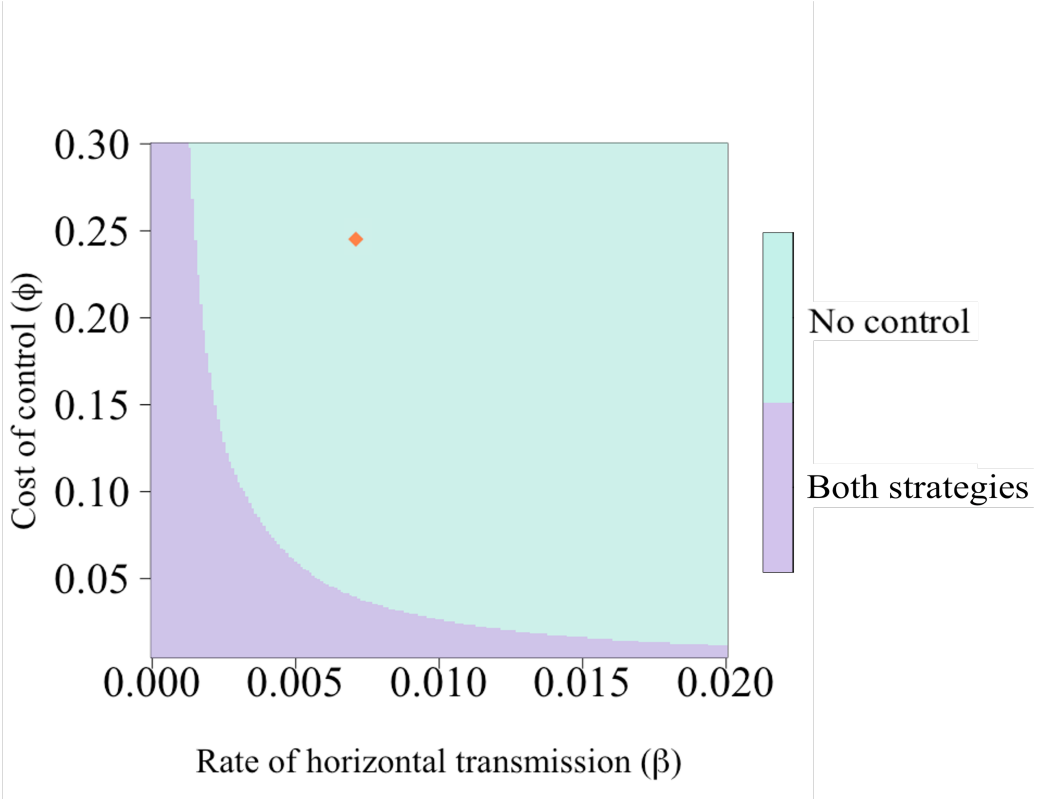

Figure 1: Possible equilibria for the "grower vs. population" model when $p = 1$. With the higher probability of vertical transmission, the "no control" equilibrium is possible for the "grower vs. population" model as there will be no non-infected, non-controlling (S_N) growers at equilibrium (Equation 56). However, now that $p = 1$, there can never be a disease-free equilibrium for this parameter set (as $R_0 > 1$).

Additionally, unlike in the "strategy vs." models, where the equilibrium is not dependent on the value of responsiveness of growers (η), for the "growers vs." models η does affect the final equilibrium values (see also Appendix 4 Figure 1).

D Appendix 4: Supplementary results

D.1 Effect of responsiveness (η) on the "grower vs." models.

Unlike in the "strategy vs." models, where the equilibrium values are not dependent on the value of responsiveness of growers, for the "growers vs." models η does affect the final equilibrium values. Appendix 4 Figure 1 shows how the proportion of controllers changes with varying values of

horizontal transmission (β) and η.

Interestingly, there is a directionality to the response. Below a certain threshold (which here is $\beta = 0.003301$ day $^{-1}$), an increase in η decreases the proportion controlling. The increase in η means that all growers who have the potential to switch strategy have a higher probability of doing so, and if η is sufficiently high then each switching term approaches 1. This means that there is a decrease in controllers, as all S_C and I_C growers will switch strategy, though only I_N growers will start controlling.

At higher values of β, the increase in η causes an increase in controllers. For these values, z_{sc} eventually falls to zero, and having higher values of η increases the rate at which this happens. As the parameter values permit an two-strategy equilibrium, but the S_C growers are no longer switching strategy, there is an increase in the number of controllers with an increase in η.

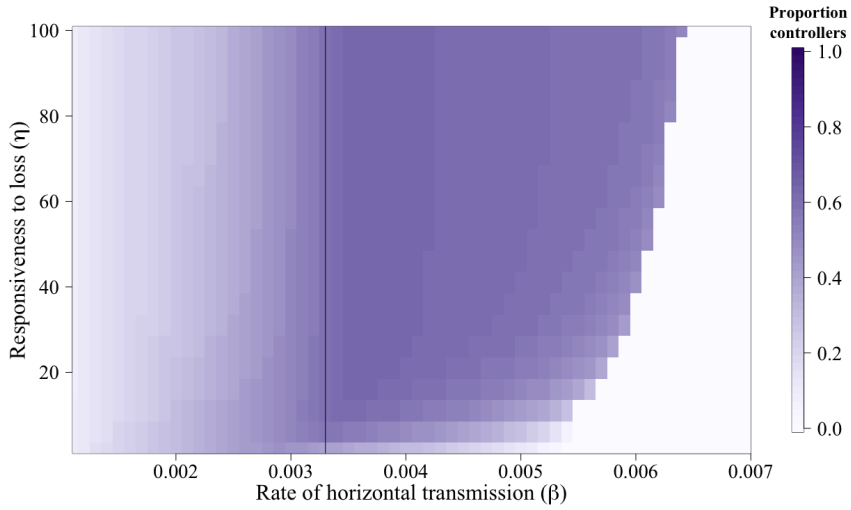

Figure 1: Effect of changes in responsiveness and horizontal transmission on the proportion of controllers at equilibrium. Below a certain threshold, there is a decrease in the proportion of controllers with an increase in η, as the higher responsiveness causes more S_C growers to switch strategy. Above this threshold, an increase in η causes an increase in controllers. The parameter values mean that S_C growers no longer change strategy, so they cannot leave the CSS. However, an increase in η means that I_N growers have a higher probability of switching into the CSS. The solid vertical line denotes where this threshold is crossed ($\beta = 0.03301$ day -1).

In the "grower vs." models, equilibrium is reached when $z_{SC}S_C + z_{IC}I_C = z_{IN}I_N$ (i.e. the flow of growers out of one strategy matches the flow of growers out of the other) (Figure 2). As the

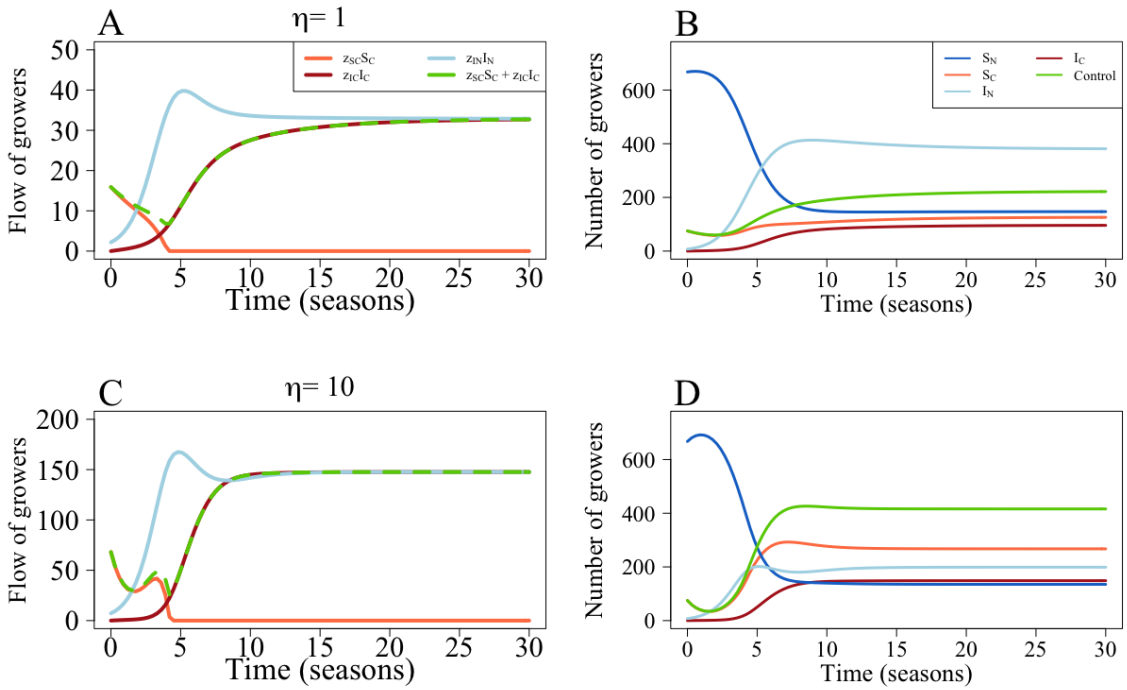

Figure 2: Effect of responsiveness (η) on the "grower vs. alternative" model. (A) The flow of growers between the non-infected controllers (S_C), infected controllers (I_C) and infected non-controllers (I_N) based on their probability of switching strategy (z_{SC}, z_{IC} and z_{IN} respectively) with $\eta = 1$. (B) Full model dynamics for $\eta = 1$. (C) The flow of growers between S_C, I_C and I_N based on their probability of switching with $\eta = 10$. Note that will lower values of η, there are fewer growers moving between strategies. (D) Full model dynamics for $\eta = 10$. Parameters are as in Table 1, except for $\beta = 0.004$ day -1, which was used to allow for an disease- and control-endemic equilibrium (Figure 2).

switching terms are in part set by the value of η, for changes in responsiveness then the values of S_C, I_C and I_N at equilibrium must change to ensure the equilibrium condition is still met.

D.2 Effect of parameters on the "strategy vs." and "grower vs. population" models.

We will now consider the three other behavioural models (the "strategy vs. population", "strategy vs. alternative" and "grower vs. population") that were not a focus of the main text. Each model formulation responded very differently to changes in parameter values. In all cases, when the rate of horizontal transmission (β) is sufficiently low, no grower should use the CSS (Appendix 4 Figure

3). At medium-to-low values of β, more growers use the CSS, though as β increased, the higher probability of infection narrows the range of costs for which a controller will consider participation in the CSS. The two "strategy vs." models - which had the same equilibria" - allowed an "all control" equilibrium at low costs of control (ϕ). This was never possible for the "grower vs. population model, as controllers managing infected fields should always consider switching strategy as they have received the "sucker's payoff".

Overall, the "strategy vs." models saw lower participation in the CSS than the "grower vs." models. The high default value of the cost of control ($\phi = 0.25$) means that growers will participate in the CSS only at very high probabilities of vertical transmission (p) and loss due to disease (L). In the "grower vs." models, however, the non-controllers with infected fields will likely have a lower payoff than the population average, and thus should have a higher probability of switching strategy for a wider range of these parameter values.

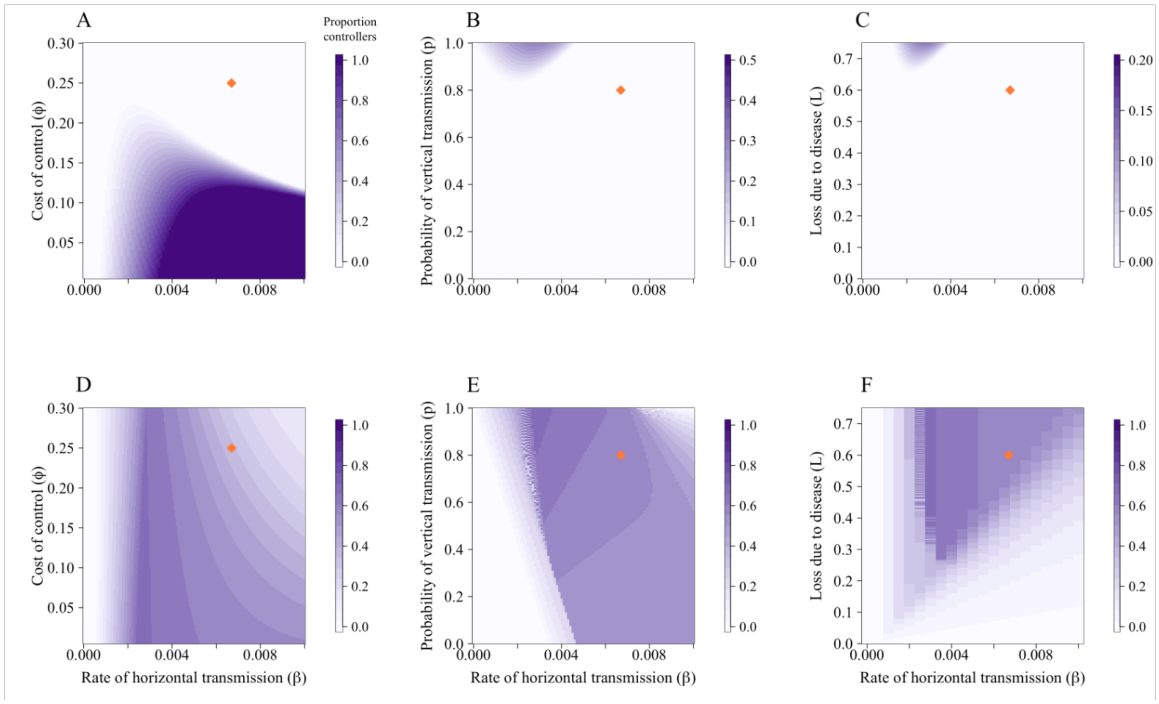

Figure 3: Effect of parameters on participation in the CSS. (A) – (C) shows results for the "strategy vs. population" and "strategy vs. alternative" model formulations; (D) – (F) are for the "grower vs. population" model. Parameters are as in Table 1 in the main text; the orange diamonds mark the default values. (A) and (D) examine the impact of changing the cost of control (ϕ) and the rate of horizontal transmission (β). At low values of β, no-one should control as the probability of infection is sufficiently low that it is not necessary. As β increases, controlling is more beneficial and the cost of control is perceived to be worthwhile. However, as infection becomes more likely, the value of control diminishes so it is not worthwhile to invest in control. In the "strategy vs." comparisons, it is possible to reach an "all control" equilibrium, though this cannot happen for the "grower vs. population" model. (B) and (C) At very high probabilities of vertical transmission (p) and loss due to disease (L), growers will participate in the CSS, though only for a narrow range of β. However, for the "grower vs. population" models, a much wider range of all parameters allowed for participation. The irregular contours in (E) and (F) are due to oscillations around the equilibrium.

D.3 Effect of subsidy on parameter scans.

Using the default parameters, there is a narrow range of parameter values for which growers should consider control at equilibrium in the "strategy vs." models (Appendix 4 Figure 4(D)). In the "grower vs." models, though there is a wider range of parameters for which growers control at equilibrium, there are still low levels of participation in the CSS (Appendix 4 Figure 4(E) - (F)). Even when the probability of vertical transmission is high, there are few growers using the CSS due to the high cost of participation. However, providing a 50% subsidy such that $\phi = 0.125$ increases both the level of participation and the range of parameter values for which growers control (Appendix 4 Figure 4).

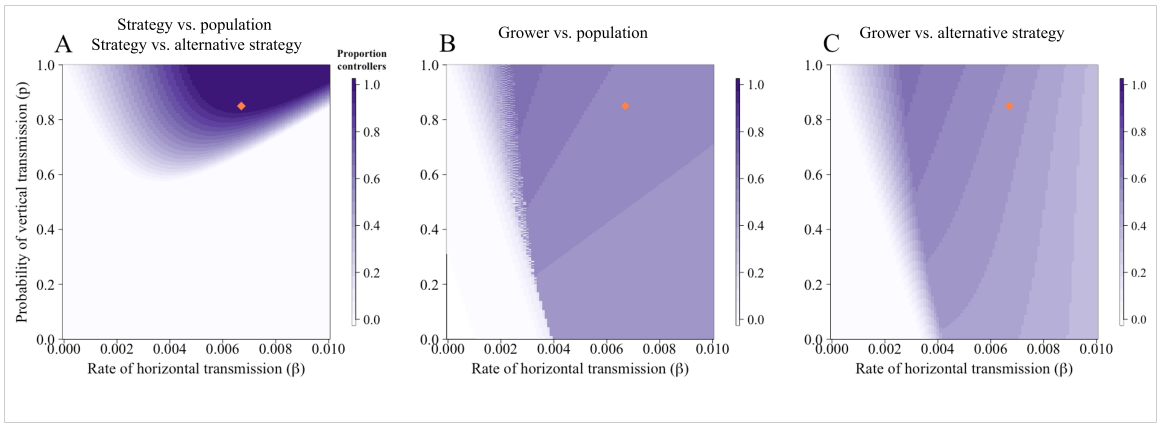

Figure 4: Effect of changes in the rate of secondary transmission (β) and probability of vertical transmission (p) on the proportion of controllers when $\phi = 0.125$. Compared to the default value of $\phi = 0.25$, a higher proportion of growers control for a broader range of parameter values.