Contributions to Nonparametric Predictive Inference for Bernoulli Data with Applications in Finance

Junbin Chen

A Thesis presented for the degree of
Doctor of Philosophy

Department of Mathematical Sciences
University of Durham
England

Jun. 2019

Dedicated to

My parents and sister

Contributions to Nonparametric Predictive Inference for Bernoulli Data with Applications in Finance

Junbin Chen

Submitted for the degree of Doctor of Philosophy
Jun. 2019

Abstract

Imprecise probability is a more general probability theory which has many advantages over precise probability theory in uncertainty quantification. Many statistical methodologies within imprecise probability framework have been developed today, one of which is nonparametric predicted inference (NPI). NPI has been developed to handle various data types and has many successful applications in different fields.

This thesis firstly further developed NPI for Bernoulli data to address two current challenging issues, the computation of imprecise expectation for a general function of multiple future stages observations and handling of imprecise Bernoulli data. To achieve the former, we introduce the concept of the mass function from Weichselberger's axiomatization of imprecise probability theory [39] and Dempster-Shafer's notion of basic probability assignment [26, 34]. Based on the concept of mass function, an algorithm to find the imprecise expectation measure for a general function of a finite random variable is proposed. We then construct mass functions of single and multiple future stages observations in NPI for Bernoulli data by its underlying latent variable representation, which leads to the applicability of the proposed algorithm in NPI for Bernoulli data. To achieve the latter, we extend the original NPI path counting method in its underlying lattice representation. This leads to the development of mass function and the imprecise probabilities of NPI for imprecise Bernoulli data. The property of NPI for imprecise Bernoulli data is illustrated with a numerical example.

Subsequently, under the binomial tree model, NPI for Bernoulli data and imprecise data are applied to asset and European options trading and NPI for Bernoulli data is applied to portfolio assessment. The performances of both applications are evaluated via simulations. The predictive nature and ability of noise recognition of NPI for precise and imprecise Bernoulli data are validated. The viability for application of NPI in portfolio assessment is confirmed.

Declaration

The work in this thesis is based on research carried out at the Department of Mathematical Sciences, Durham University, England. No part of this thesis has been submitted elsewhere for any other degree or qualification and it is all my own work unless referenced to the contrary in the text.

Part of Chapter 4 and 5 was presented at 7th International Conference of the Financial Engineering and Banking Society Glasgow in June 2017 and submitted to "Journal of The Operational Research Society" [7]

Acknowledgements

My journey in Ph.D. study is not as smooth as it was expected. I quitted once for what I believed a have-to-make decision, and I came back to recontinue with a broken heart.

For this, I would like to firstly thank Prof. Frank Coolen enormously. Not only because he constantly gives me expert advice and generously shares his profound insights with me in my research problems, but also because he patiently listened to me and gave personal suggestions I was low and took me back unconditionally when I decided to recontinue my Ph.D. study.

Equally, I would like to thank my parents and my sister who has accompanied me and given me various support through my hard time. Although those difficult time of mine has been far away, I still remember, and I am genuinely appreciated and deeply indebted to any of you.

Also, I would like to thank Dr.Tahani Coolen-Maturi for her useful comments and suggestions during the process of writing my thesis.

I am an atheist, but sometimes I do believe the existence of the gap years during my Ph.D. study is for a beautiful reason — to meet many pure and joyful minds like my friends here. So finally and importantly, I would to thank my friends Calum, Clare, Dan, Daniel, Hugo, James, Joe, Pla, Rob, Themis, Tim, Ting in the mathematics department, some for great banters, some for witty conversations, some for valuable discussions, some for mindful table gaming, some for enjoyable live music, some for delicious dinners and some for good stories.

Contents

Chapter 1

Introduction

In this chapter, the background of imprecise probability is briefly recalled within which the motivations of this thesis are highlighted. Subsequently, the outline of content is presented which elaborate logical structure of this thesis.

1.1 Motivations

Imprecise probability or sometimes called interval probability is a more general framework of probability theory. Its development could date back to 1854, by Boole [6]. From the time, most contributions had been made to the reconciliation between theories of logic and probability. Later, the notion of imprecise probability has been advocated by several authors including Peter Walley, Kurt Weichselberger, et al. [36, 38–40].

But, why do we need imprecise probability, a more general probability theory which quantifies uncertainty by a set of probability measure instead of one single probability measure? There are many reasons for this and some of them from the data perspective are illustrated below.

In the real application, in order to train the model properly, sufficient data must be gathered. However, to gather enough data is not always possible in practice. In this situation, precise probability usually falls short of applicability, as a single probability measure can hardly be deduced accurately due to lack of data. Imprecise probability, on the other hand, provides a more viable resolution to this situation.

Instead of using one single probability measure, imprecise probability uses a set of probability measures which allows a degree of imprecision in the inference.

Granted that one can gather enough data to train the model, noises within data are inevitable. When noise is contained in the data, modelling the uncertainty with a single probability may not be justified, as the used single probability is likely different from the true underlying probability. On the other hand, a set probability measure is more likely to cover the true underlying distribution. Imprecise probability again, in this case, is a more appropriate theory to be used.

Hence, imprecise probability seems to be a more applicable theory to model uncertainties in reality as lack of data and noise contained in the data constantly happen in the real practices.

Nowadays many imprecise probability methodologies have been developed, one of which is nonparametric predictive inference (NPI) developed by Coolen [8,13,18]. It has been developed to handle different data types and has many successful applications in the field of engineering reliability. The existing researches have shown NPI always give consistent results. However, the current development of NPI for Bernoulli data is facing two unsolved issues—the computation of imprecise expectation for a general function of multiple future stages observations and handling of imprecise Bernoulli data. Addressing these two issues is then the first motivation of this thesis. Also, modeling financial uncertainty using imprecise probability appears to have more advantages than its precise probability [35] and little effort has been dedicated to the NPI's application in finance so far. NPI for Bernoulli data may not be a suitable method to model a sequence of future Bernoulli events which is not close to identical distributed due to its positive learning from historical data. It is, however, a suitable method to model a sequence of future Bernoulli events which are approximately identically distributed but not necessarily independent. When considering a certain asset over a short period of time on the binomial tree model, one could assume the market participants over this time period are approximately homogeneous. Thus the asset price upward or downward movement in each time stage is approximately identically distributed and is suitable modeled by NPI for Bernoulli data. Hence, the second motivation of this thesis is to apply NPI for

Bernoulli data in finance trading.

1.2 Outline of thesis

Chapter 2 presents preliminary material in this thesis which includes the basic framework language used in this thesis, review of NPI, relevant financial concepts and objects of interest in this thesis. It begins with a introduction of a set of mass function based imprecise probability definitions. The idea of mass function comes from Weichselberger's axiomatization of imprecise probability [39] and Dempster-Shafer's notion of basic probability assignment [26,34]. The introduced definitions will serve as the basic framework language in Chapter 3. Afterward, the imprecise probability methodology — Nonparametric predictive inference (NPI) is introduced within which the current development of NPI for Bernoulli data is reviewed in detail, and two of its current challenging issues are identified. These are essentially the motivations of Chapter 3. In the end, with reasonable assumptions, the financial objects for later NPI application are defined, relevant financial concepts are introduced, and some financial terminologies are explained which provides necessary information for one who is less familiar in finance.

The aim in Chapter 3 is to address two challenging issues in NPI for Bernoulli data identified in Chapter 2. To achieve this, a general algorithm to find imprecise expectation measures for a general function of a finite random variable in an imprecise probability space is firstly presented, which provides a tool to address the first issue. Second, in order to enable the usage of the presented algorithm in NPI for Bernoulli data, the mass function of NPI is constructed using its latent variable representation. By using a mapping between NPI imprecise probability and path counting within a lattice, the constructed mass function is shown to produce the same imprecise probability as presented in Chapter 2. The consistence of the constructed mass function is also proved. With the presented algorithm and constructed mass function, a complete example of how to use the algorithm to construct imprecise expectation measures for a general function of future observations in NPI for Bernoulli data is presented. Finally, by extending NPI path counting method in its underlying lattice representation, NPI for imprecise Bernoulli data is developed which addresses the second issue of NPI for Bernoulli data.

In Chapter 4, under the binomial tree model, NPI for Bernoulli data and imprecise Bernoulli data are applied in financial asset trading in a prescribed scenario. Two trading routes with different trading primary objectives are proposed and computer simulation is conducted to evaluate the performance of the trading routes under the different market conditions and data imprecision. The result shows that the proposed trading routes for asset are able to execute correct action according to the situation, has good predictivity and noise recognition.

In Chapter 5, under the binomial tree model, NPI for Bernoulli data and imprecise Bernoulli data are applied in financial European call option and European put option trading in two separate prescribed scenarios. Trading routes with different trading primary objectives for both call options and put options are proposed. Computer simulation is conducted to evaluate the performance of the proposed trading routes under different market conditions and data imprecision. The simulation result confirms that the proposed NPI trading routes have good predictivity, quick learning property, and moderate noise resistance.

In Chapter 6, under the binomial tree model, NPI for Bernoulli data is applied in financial portfolio assessment. Computer simulation is conducted to evaluate the performance of NPI assessment method proposed. The viability for application of NPI in portfolio assessment is confirmed.

In Chapter 7, a general conclusion of the thesis is drawn. Some of the potential future extensions of the research presented are suggested.

Chapter 2

Preliminaries

In this chapter, using the concept of the mass function from Weichselberger's axiomatization of imprecise probability [39] and Dempster-Shafer's notion of basic probability assignment [26, 34], a set of mass function based imprecise probability definitions is firstly introduced, which serves as a basic framework in Chapter 3. After that, the imprecise probability methodology — Nonparametric predictive inference (NPI) is introduced. Specifically, the current development of NPI for Bernoulli data is reviewed in detail in which two of its current challenging issues are identified. Next, with reasonable assumptions, the financial objects for later NPI application are mathematically defined. Also, relevant financial concepts are introduced. Finally, some financial terminologies are explained which provides necessary information for one who is less familiar in finance.

2.1 Imprecise probability definitions

In this section, a set of mass function based imprecise probability definitions is introduced. It is should be noted that throughout this thesis all sample spaces Ω considered are countable.

Definition 2.1.1 (Precise probability space $\mathscr{K}$)

Given a sample-space Ω, a sigma algebra $\mathscr{A}$ of a collection of events in Ω, and a set function $p : \mathscr{A} \longrightarrow [0, 1]$, the triple $\mathscr{K} = [\Omega, \mathscr{A}, p]$ is called a precise probability space if p satisfies Kolmogorov axiom (I–III):

I: $p(\theta) \geqslant 0 \quad \forall \theta \in \mathscr{A}$

II: $p(\Omega) = 1$

III: If $\theta_i \in \mathscr{A}$ for $i \in \mathbb{N}$ and $\theta_i \cap \theta_j = \emptyset$ for $i \neq j$, then $p(\cup_{i\in\mathbb{N}}\theta_i) = \sum_{i\in\mathbb{N}} p(\theta_i)$

p is called a precise probability for $(\Omega; \mathscr{A})$

Definition 2.1.2 (A set P of all precise probabilities for $[\Omega, \mathscr{A}]$)

Given a measurable space $(\Omega; \mathscr{A})$, we denote all precise probabilities of this space as P.

$P = \{p|\ p \text{ satisfies Kolmogorov axiom (I-III) in } (\Omega, \mathscr{A})\}$

Definition 2.1.3 (Imprecise probability space $\mathscr{I}$)

Given a sample-space Ω and a sigma algebra $\mathscr{A}$ of a collection of events in Ω, a set function $m(\cdot)$ mapping from elements in $\mathscr{A}$ to $[0,1]$, $m(\cdot) : \mathscr{A} \longrightarrow [0,1]$.

The triple $\mathscr{I} = [\Omega, \mathscr{A}, m(\cdot)]$ is a imprecise probability space $\mathscr{I}$ if $m(\cdot)$ satisfies the following conditions:

I: $m(\emptyset) = 0$; $m(\epsilon) \geq 0$, $\forall \epsilon \in \mathscr{A}$

II: $\sum\limits_{\epsilon \in \mathscr{A}} m(\epsilon) = 1$

$m(\cdot)$ is called a mass function for $[\Omega, \mathscr{A}]$

Given one imprecise probability space $\mathscr{I} = [\Omega, \mathscr{A}, m(\cdot)]$ defined as above. The corresponding upper probability $\overline{p}$ and lower probability $\underline{p}$ based on the mass function $m(\cdot)$ of a event $\mu \in \mathscr{A}$ are defined as:

$$\overline{p}(\mu) = \sum_{\substack{\epsilon \in \mathscr{A} \\ \epsilon \cap \mu \neq \emptyset}} m(\epsilon) \quad \text{and} \quad \underline{p}(\mu) = \sum_{\substack{\epsilon \in \mathscr{A} \\ \epsilon \subset \mu}} m(\epsilon)$$

Conjugacy property of the upper and lower probability

By Definition 2.1.3, for any $[\Omega, \mathscr{A}, m(\cdot)]$, there is a conjugacy property between $\overline{p}(\cdot)$ and $\underline{p}(\cdot)$ as follows.

For an event $\mu \in \mathscr{A}$, let μ^c denote the complement of μ. $\mu^c \cup \mu = \Omega$, then:

$\overline{p}(\mu^c) + \underline{p}(\mu) = 1$

To show this, first, let us prove two propositions.

Proposition 2.1.1

$\{\epsilon | \epsilon \in \mathscr{A}, \epsilon \cap \mu^c \neq \emptyset\} \cap \{\epsilon | \epsilon \in \mathscr{A}, \epsilon \subset \mu\} = \emptyset$

Proposition 2.1.2

$\{\epsilon | \epsilon \in \mathscr{A}, \epsilon \cap \mu^c \neq \emptyset\} \cup \{\epsilon | \epsilon \in \mathscr{A}, \epsilon \subset \mu\} = \{\epsilon | \epsilon \in \mathscr{A}\} = \mathscr{A}$

Proof:

For Proposition 2.1.1. If $\theta \in \{\epsilon | \epsilon \in \mathscr{A}, \epsilon \subset \mu\}$, then $\theta \cap \mu^c = \emptyset$. Thus $\theta \notin \{\epsilon | \epsilon \in \mathscr{A}, \epsilon \cap \mu^c \neq \emptyset\}$

For Proposition 2.1.2. If $\theta = \{\epsilon | \epsilon \in \mathscr{A}, \epsilon \cap \mu^c \neq \emptyset\}$, then $\theta^c = \{\epsilon | \epsilon \in \mathscr{A}, \epsilon \cap \mu^c = \emptyset\} = \{\epsilon | \epsilon \in \mathscr{A}, \epsilon \subset \mu\}$

Using Propositions 2.1.1 and 2.1.2, one then has:

$$\sum_{\epsilon \in \mathscr{A}} m(\epsilon) = \sum_{\substack{\epsilon \in \mathscr{A} \\ \epsilon \cap \mu^c \neq \emptyset}} m(\epsilon) + \sum_{\substack{\epsilon \in \mathscr{A} \\ \epsilon \subset \mu}} m(\epsilon) \qquad \forall \mu \in \mathscr{A} \tag{2.1.1}$$

$$1 = \overline{p}(\mu^c) + \underline{p}(\mu) \qquad \text{By Definition 2.1.3} \tag{2.1.2}$$

Thus one can also have the following:

1. $\underline{p}(\emptyset) = \sum_{\substack{\epsilon \in \mathscr{A} \\ \epsilon \subset \emptyset}} m(\epsilon) = 0$
2. $\overline{p}(\Omega) = 1 - \underline{p}(\emptyset) = 1$ by Equality 2.1.2
3. $\underline{p}(\Omega) = \sum_{\substack{\epsilon \in \mathscr{A} \\ \epsilon \subset \Omega}} m(\epsilon) = 1$
4. $\overline{p}(\emptyset) = 1 - \underline{p}(\Omega) = 0$ by Equality 2.1.2

Definition 2.1.4 (Atom event)

If an event $\epsilon = \{Q\} \in \mathscr{A}$ contains only one element Q in the sample space Ω ($Q \in \Omega$), we call this event an atom event.

Interpretation of the mass function on non atom event

Given $[\Omega, \mathscr{A}, m(\cdot)]$, the value that a mass function assigns to a non atom event could be understood as the shared mass or uncertain mass between the atoms. For example, for event $E = \{Q_1, Q_2, Q_3\}$ where $Q_1, Q_2, Q_3 \in \Omega$, the mass value $m(E)$ can be understood as the shared mass between Q_1, Q_2, Q_3. In other words, the mass value $m(E)$ can be assigned to event Q_1 or Q_2 or Q_3, but it does not necessarily need to be assigned to Q_1 or Q_2 or Q_3. When one takes the upper probability of $\{Q_1\}$, $\overline{p}(\{Q_1\}) = \sum_{\substack{\epsilon \in \mathscr{A} \\ \epsilon \cap Q_1 \neq \emptyset}} m(\epsilon)$, which is an optimistic probability evaluation of Q_1, the mass value $m(E)$ is included. In contrast, when one takes the lower probability of $\{Q_1\}$, $\underline{p}(\{Q_1\}) = \sum_{\substack{\epsilon\in\mathscr{A}\\ \epsilon\subset Q_1}} m(\epsilon)$, which is a conservative probability evaluation of Q_1, the mass value $m(E)$ is excluded.

In the case where all mass values for non atom events are zero, the imprecise probability space $[\Omega, \mathscr{A}, m(\cdot)]$ with finite sample space Ω become a precise probability space $[\Omega, \mathscr{A}, p(\cdot)]$. That is, $p(\{Q\}) = m(\{Q\})$, $\forall Q \in \Omega$.

Definition 2.1.5 (Consistence of a sequence of mass functions)
Given an index set I and $\mathscr{I}_i = [\Omega, \mathscr{A}_i, m_i(\cdot)]$ a sequence of mass functions $m_i(\cdot)$ defined on different event spaces $\mathscr{A}_i$ with respect to same sample space Ω, $i \in I$. The sequence of mass functions $\{m_i(\cdot)\}$ is said to be consistently defined or consistent if $\forall \epsilon_j \in \mathscr{A}_j$, $\forall \epsilon_k \in \mathscr{A}_k$, $j \neq k$, $\epsilon_j \subset \epsilon_k$, then $\underline{p}_j(\epsilon_j) \leq \underline{p}_k(\epsilon_k)$ where $\underline{p}_j$ and $\underline{p}_k$ is the lower probability induced by $m_j(\cdot)$ and $m_k(\cdot)$ respectively.

Definition 2.1.6 (A subset P_m of all precise probabilities P induced by a $m(\cdot)$)
Given a measurable space $(\Omega; \mathscr{A})$, the set P of all precise probabilities on this space and a mass function $m(\cdot)$ on this space, one can induce a subset P_m of P by the mass function $m(\cdot)$. P_m is called a credal set or structure in some literature.

$$P_m = \{p(\cdot) | p(\cdot) \in P, \sum_{\substack{\epsilon\in\mathscr{A}\\ \epsilon\subset\cdot}} m(\epsilon) \leq p(\cdot) \leq \sum_{\substack{\epsilon\in\mathscr{A}\\ \epsilon\cap\cdot\neq\emptyset}} m(\epsilon)\}$$

Thus within $[\Omega, \mathscr{A}, P_m]$, one has
$\inf_{p(\cdot)\in P_m} p(\theta) = \underline{p}(\theta)$ and $\sup_{p(\cdot)\in P_m} p(\theta) = \overline{p}(\theta) \quad \forall\theta \in \mathscr{A}$

By using imprecise probability, one now can use a single mass function and work on the induced probabilities P_m instead of using a single probability in the application. By doing this, the model can be more robust than its precise probability counterpart as a set of probabilities is more likely to cover the true underlying probability of the uncertainties. Also, since gathering perfect information is not always possible in practice, imprecise model would be a more appropriate model to reflect the lack of perfect information.

Definition 2.1.7 (Discrete Random variable X)
A discrete random variable is a function $X : \Omega \longrightarrow F$ where F is a countable ordered

field and $X^{-1}(x) \in \mathscr{A}\ \forall\ x \in F$

Definition 2.1.8 (Imprecise expectation of a discrete random variable)
Given an imprecise probability space $[\Omega, \mathscr{A}, m(\cdot)]$, a discrete random variable is a function $X : \Omega \longrightarrow F$. We define the lower expectation $\underline{E}$ and the upper expectation $\overline{E}$ of X as:

$$\underline{E}(X) = \inf_{p(\cdot)\in P_m} \sum_{\omega\in\Omega} X(\omega)p(\omega) \tag{2.1.3}$$

$$\overline{E}(X) = \sup_{p(\cdot)\in P_m} \sum_{\omega\in\Omega} X(\omega)p(\omega) \tag{2.1.4}$$

Definition 2.1.9 (The lower and upper expectation measure of X)
Given an imprecise probability space $[\Omega, \mathscr{A}, m(\cdot)]$, a discrete random variable is a function $X : \Omega \longrightarrow F$, We then define the lower expectation measure $p_{\underline{E}(X)}(\cdot)$ and and upper expectation measure $p_{\overline{E}(X)}(\cdot)$ as:

$$p_{\underline{E}(X)}(\cdot) = \operatorname*{argmin}_{p(\cdot)\in P_m} \sum_{\omega\in\Omega} X(\omega)p(\omega) \tag{2.1.5}$$

$$p_{\overline{E}(X)}(\cdot) = \operatorname*{argmax}_{p(\cdot)\in P_m} \sum_{\omega\in\Omega} X(\omega)p(\omega) \tag{2.1.6}$$

To compute the imprecise expectation of a random variable or a function of a random variable $f(X)$, one needs to find a way construct $p_{\underline{E}(X)}(\cdot)$ and $p_{\overline{E}(X)}(\cdot)$ or $p_{\underline{E}(f(X))}(\cdot)$ and $p_{\overline{E}(f(X))}(\cdot)$. Based on the above definitions of imprecise probability, an algorithm for the construction of imprecise expectation measures for a general function of a finite random variable is presented in Chapter 3.

Definition 2.1.10 (Product space of independent spaces)
A finite sequence of imprecise probability spaces $[\Omega_i, \mathscr{A}_i, m_i(\cdot)]_{i=1}^{i=n}$ are mutually independent if $\Omega_i \cap \Omega_j = \emptyset$ when $i \neq j$.
$[\Omega, \mathscr{A}, m(\cdot)]$ is defined as the product space of n independent imprecise probability space if $\Omega = \Omega_1 \times \Omega_2 \times ... \times \Omega_n$, $\mathscr{A} = \mathscr{A}_1 \times \mathscr{A}_2 \times ... \times \mathscr{A}_n$, and $m(\cdot) = \prod_{i=1}^{i=n} m_i(\cdot)$

Given n finite random variables X_i on different imprecise probability spaces $[\Omega_i, \mathscr{A}_i, m_i(\cdot)]_{i=1}^{i=n}$, $X_i : \Omega_i \longrightarrow F_i$. $i \in \{1, 2, ..n\}$, then on the product space

$[\Omega, \mathscr{A}, m(\cdot)]$ as defined in Definition 2.1.10, one has:

$$\underline{p}(\cap_{i=1}^{i=n} X_i \in W_i) = \prod_{i=1}^{n} \underline{p}_i(X_i \in W_i) \quad \forall W_i \subset F_i \tag{2.1.7}$$

$$\overline{p}(\cap_{i=1}^{i=n} X_i \in W_i) = \prod_{i=1}^{n} \overline{p}_i(X_i \in W_i) \quad \forall W_i \subset F_i \tag{2.1.8}$$

Proof:

$$\begin{aligned}
\underline{p}(\cap_{i=1}^{i=n} X_i \in W_i) &= \sum_{\substack{\forall i \in \mathbb{N}_1^n, \forall W_i^j \\ W_i^j \subset W_i \text{ and } X_i^{-1}(W_i^j) \in \mathscr{A}_i}} m(\cap_{i=1}^{i=n} X_i \in W_i^j) \\
&= \sum_{\substack{\forall i \in \mathbb{N}_1^n, \forall W_i^j \\ W_i^j \subset W_i \text{ and } X_i^{-1}(W_i^j) \in \mathscr{A}_i}} \prod_{i=1}^{n} m_i(X_i \in W_i^j) \\
&= \prod_{i=1}^{n} \sum_{W_i^j \subset W_i \text{ and } X_i^{-1}(W_i^j) \in \mathscr{A}_i} m_i(X_i \in W_i^j) \\
&= \prod_{i=1}^{n} \underline{p}_i(X_i \in W_i)
\end{aligned}$$

The first line of the proof used Definition 2.1.3, the second line used Definition 2.1.10, the third line used the fact interchanging $\sum$ and $\prod$ have the same mass value cumulation in the equation, the fourth line used Definition 2.1.3. Similarly, one has:

$$\begin{aligned}
\overline{p}(\cap_{i=1}^{i=n} X_i \in W_i) &= \sum_{\substack{\forall i \in \mathbb{N}_1^n, \forall W_i^j \\ W_i^j \cap W_i \neq \emptyset \text{ and } X_i^{-1}(W_i^j) \in \mathscr{A}_i}} m(\cap_{i=1}^{i=n} X_i \in W_i^j) \\
&= \sum_{\substack{\forall i \in \mathbb{N}_1^n, \forall W_i^j \\ W_i^j \cap W_i \neq \emptyset \text{ and } X_i^{-1}(W_i^j) \in \mathscr{A}_i}} \prod_{i=1}^{n} m_i(X_i \in W_i^j) \\
&= \prod_{i=1}^{n} \sum_{W_i^j \cap W_i \neq \emptyset \text{ and } X_i^{-1}(W_i^j) \in \mathscr{A}_i} m_i(X_i \in W_i^j) \\
&= \prod_{i=1}^{n} \overline{p}_i(X_i \in W_i)
\end{aligned}$$

Therefore, the structure P_m of $[\Omega, \mathscr{A}, m(\cdot)]$ is:

$$\begin{aligned} P_m &= \{p(\cdot)|p(\cdot) \in P,\ \underline{p}_m(\cdot) \leq p(\cdot) \leq \overline{p}_m(\cdot)\} \\ &= \{p(\cdot)|p(\cdot) \in P,\ \prod_{i=1}^{n} \underline{p}_{m_i}(\cdot) \leq p(\cdot) \leq \prod_{i=1}^{n} \overline{p}_{m_i}(\cdot)\} \end{aligned}$$

Moreover, if we denote P_{m_i} as structure of $[\Omega_i, \mathscr{A}_i, m_i(\cdot)]$. P_i as the set of all the precise probability measure for $[\Omega_i, \mathscr{A}_i]$, then:

$$P_{m_i} = \{p_i(\cdot)|p_i(\cdot) \in P_i,\ \underline{p}_{m_i}(\cdot) \leq p_i(\cdot) \leq \overline{p}_{m_i}(\cdot)\}$$

One also has:

$$P_m = \{p(\cdot)|p(\cdot) = \prod_{i=1}^{i=n} p_i(\cdot)\ ,\ p_i(\cdot) \in P_{m_i}\} \tag{2.1.9}$$

For imprecise expectation of sum of independent random variables as above, one has:

$$\underline{E}(\sum_{i=1}^{i=n} X_i) = \sum_{i=1}^{i=n} \underline{E}(X_i) \tag{2.1.10}$$

$$\overline{E}(\sum_{i=1}^{i=n} X_i) = \sum_{i=1}^{i=n} \overline{E}(X_i) \tag{2.1.11}$$

Proof:

$$\begin{aligned}
\underline{E}(\sum_{i=1}^{i=n} X_i) &= \inf_{p(\cdot)\in P_m} \sum_{\omega\in\Omega}(\sum_{i=1}^{i=n} X_i(\omega_i))p(\omega) \quad \text{by Definition 2.1.8} \\
&= \inf_{p(\cdot)\in P_m} \sum_{\omega\in\Omega}(\sum_{i=1}^{i=n} X_i(\omega_i))\prod_{i=1}^{i=n} p_i(\omega_i) \quad \text{by Equality 2.1.9} \\
&= \inf_{p(\cdot)\in P_m} \sum_{\substack{\forall i\in\mathbb{N}_1^n \\ \omega_i\in\Omega_i}} (\sum_{i=1}^{i=n} X_i(\omega_i)\prod_{i=1}^{i=n} p_i(\omega_i)) \\
&= \inf_{p(\cdot)\in P_m} (\sum_{i=1}^{i=n} \sum_{\substack{\forall i\in\mathbb{N}_1^n \\ \omega_i\in\Omega_i}} X_i(\omega_i)\prod_{i=1}^{i=n} p_i(\omega_i)) \\
&= \inf_{p_i(\cdot)\in P_{m_i}} (\sum_{i=1}^{i=n} \sum_{\omega_i\in\Omega_i} X_i(\omega_i)p_i(\omega_i)) \\
&= \sum_{i=1}^{i=n} \inf_{p_i(\cdot)\in P_{m_i}} \sum_{\omega_i\in\Omega_i} X_i(\omega_i)p_i(\omega_i) \\
&= \sum_{i=1}^{i=n} \underline{E}(X_i)
\end{aligned}$$

To prove Equality 2.1.11, one only need to change inf to sup in the above argument. It is hence omitted here.

Sometimes one maybe also interested in the lower probability of the event $\sum_{i=1}^{i=n} X_i > \lambda$ where λ is real value. Let e_i denote one of possible value that finite random variable X_i could take, in other words $e_i \in F_i$, then one has:

$$\begin{aligned}
\underline{p}(\sum_{i=1}^{i=n} X_i > \lambda) &= \inf_{p\in P_m} (\sum_{i=1}^{i=n} X_i > \lambda) \quad \text{by Definition 2.1.6} \\
&= \inf_{\substack{\forall i\in\mathbb{N}_1^n \\ p_i\in P_{m_i}}} \sum_{e_2,e_3,\ldots e_n} p_1(X_1 > \lambda - \sum_{i=2}^{i=n} e_i)\prod_{i=2}^{i=n} p_i(X_i = e_i) \quad \text{by Equality 2.1.9} \qquad (2.1.12) \\
&= \sum_{e_2,e_3,\ldots e_n} p_{\underline{E}(X_1)}(X_1 > \lambda - \sum_{i=2}^{i=n} e_i)\prod_{i=2}^{i=n} p_{\underline{E}(X_i)}(X_i = e_i) \qquad (2.1.13)
\end{aligned}$$

The last line comes from the fact that the lower expectation measure $p_{\underline{E}(X_i)}$ assign as least mass value as possible to the greater value X_i could take. Since each X_i is independent, each p_i can be chosen from each structure P_{m_i} independently in Equation 2.1.12. For $\sum\limits_{i=1}^{i=n} X_i > \lambda$, in Equation 2.1.12, the less value e_i in each X_i takes, the more likely "$p_1(X_1 > \lambda - \sum\limits_{i=2}^{i=n} e_i)$" will result in zero value in the formation of the product "$p_1(X_1 > \lambda - \sum\limits_{i=2}^{i=n} e_i) \prod\limits_{i=2}^{i=n} p_i(X_i = e_i)$", also the greater value e_i in each X_i takes, the more frequent that value $X_i = e_i$ will be used in formation of the non-zero product "$p_1(X_1 > \lambda - \sum\limits_{i=2}^{i=n} e_i) \prod\limits_{i=2}^{i=n} p_i(X_i = e_i)$" within the summation. So to minimize valuation of the expression "$\sum\limits_{e_2,e_3,\dots e_n} p_1(X_1 > \lambda - \sum\limits_{i=2}^{i=n} e_i) \prod\limits_{i=2}^{i=n} p_i(X_i = e_i)$" in Equation 2.1.12, one should allocate the least possible mass to greater value e_i each X_i could take, which, in essence is taking lower expectation measure for each independent random variable X_i.

2.2 Nonparametric predictive inference

Nonparametric predictive inference (NPI) is a imprecise probability methodology developed by Coolen [8, 13, 18]. It is a low structure statistical methodology based on Hill's $A_{(n)}$ assumption [29]. When no prior knowledge of the problem is known, NPI is a suitable method as it requires minimal modeling assumption. NPI has also shown stronger consistency than other conventional methods [19–21] in empirical study and no contradiction has been found in the inference it produced so far.

Based on $A_{(n)}$ assumption, with latent variables representable of historical data, NPI has been developed for Bernoulli data [8], real-valued data [4], data including right-censored observations [9] and multinomial data [11, 17]. It now currently has many successful applications in engineering reliability [1, 12, 14, 19, 22]. The existing researches have shown that NPI has good statistical properties and gives reliable predictive results. It also has recently been applied to the field of finance. [5, 7, 27] Yet more effort for its application in finance is still demanding.

As one of the attempts in this thesis is to further develop NPI for Bernoulli data, the Hill's $A_{(n)}$ assumption is firstly introduced below and based upon that, the current development of NPI for Bernoulli data is reviewed within which two current challenging issues are identified

2.2.1 Hill's assumption $A_{(n)}$

As previously mentioned, NPI is based on Hill's assumption $A_{(n)}$. This assumption is suitable for situations where no probability distribution regarding a future random quantity is assumed. The Hill's assumption $A_{(n)}$ is stated as follows:

Given n exchangeable real-valued observations $y_{-n+1}, y_{-n+2}, ..., y_0$ with order statistics $y_{(1)} < y_{(2)} < .... < y_{(n)}$. We define $y_{(0)} = -\infty$, $y_{(n+1)} = \infty$ and assume $p(y_i = y_j) = 0$ for $i \neq j$. Then $y_{-n+1}, y_{-n+2}, ..., y_0$ divide the real-line into $n+1$ intervals $I_g = (y_{(g-1)}, y_{(g)})$ for $g = 1, 2, ..., n+1$. The assumption $A_{(n)}$ states that a future random quantity Y_t, $t \in \mathbb{N}^+$ will fall equally likely into each interval. (See Figure 2.1).

$$p(Y_t \in I_g) = \frac{1}{n+1} \quad \text{for} \quad g = 1, 2...n+1; \;\; t \in \mathbb{N}^+$$

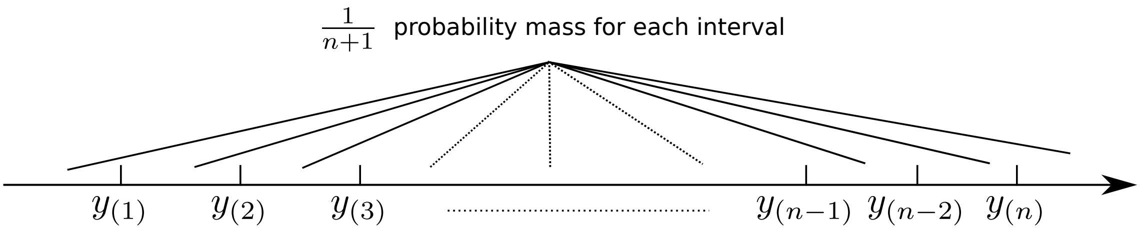

Figure 2.1: $A_{(n)}$ assumption

2.2.2 NPI for Bernoulli data

Based on the $A_{(n)}$ assumption, Coolen in 1998 developed NPI for Bernoulli data using an underlying latent variable representation [8] which will be demonstrated below.

Given n exchangeable Bernoulli observations $\{o_i\}_{i=-n+1}^{0}$, $o_i \in \{B, B^c\}$, where B and B^c represent two possible outcomes on one single observation, one then has a set $D_{(n)} = \{x_i\}_{i=-n+1}^{0}$ of Bernoulli data with $x_i \in \{0, 1\}$, $x_i(o_i) = \mathbb{1}_{\{B\}}(o_i)$. By assuming a latent threshold variable ℓ and a sequence of latent real values y_i corresponding to each observation x_i with order statistics $y_{(1)} < y_{(2)} < .... < y_{(n)}$ such that for all the data $x_i = 1$ if and only if $y_i < \ell$, $x_i = 0$ if and only if $y_i > \ell$. If one calls B a success of the event, then the number of successes in the data is $j = |\{i : x_i = 1\}| = |\{i : y_i < \ell\}|$. Since the sufficient statistics in NPI for Bernoulli data is n and j, (n, j) will be used subsequently to represented $D_{(n)}$.

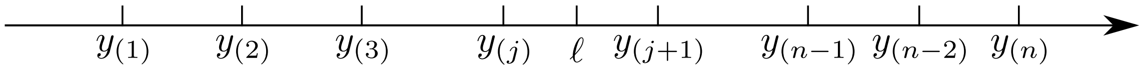

Figure 2.2: NPI for Bernoulli data underlying representation

Now by the $A_{(n)}$ assumption, given data (n, j), NPI for Bernoulli data define the lower probability $\underline{p}_{(n,j)}$ of future observation at t-th stage observing "B" (equivalently $X_t \in \{1\}$) as value of all probability mass in the interval that must be less than ℓ, and define the upper probability $\overline{p}_{(n,j)}$ of future observation at t-th stage observing "B" as value of all probability mass in the interval that could be less than "ℓ". In other words:

$$\underline{p}_{(n,j)}(X_t \in \{1\}) = |\{g : I_g \subset (-\infty, \ell)\}| = \tfrac{j}{n+1} \tag{2.2.14}$$

$$\overline{p}_{(n,j)}(X_t \in \{1\}) = |\{g : I_g \cap (-\infty, \ell) \neq \emptyset\}| = \tfrac{j+1}{n+1} \tag{2.2.15}$$

And with the same principle, for future observation at t-th stage observing "B^c" (equivalently $X_t \in \{0\}$). The NPI imprecise probability is:

$$\underline{p}_{(n,j)}(X_t \in \{0\}) = |\{g : I_g \subset (\ell, +\infty)\}| = \tfrac{n-j}{n+1} \tag{2.2.16}$$

$$\overline{p}_{(n,j)}(X_t \in \{0\}) = |\{g : I_g \cap (\ell, +\infty) \neq \emptyset\}| = \tfrac{n-j+1}{n+1} \tag{2.2.17}$$

By assuming $A_{(n)}$ up to $A_{(n+T-1)}$, NPI for Bernoulli data further define the imprecise probability for the number of observations of "B" within any future T stages. Mathematically, denote the number of observations of "B" within any future T stages as S_T and denote $\mathbb{N}^+$ as the set of all positive natural number then:

$$S_T = \sum_{t \in W} X_t \quad \text{where} \quad W \subset \mathbb{N}^+ \quad \text{and} \quad |W| = T$$

There are $\binom{n+T}{n}$ ways to distribute T future observations on the underlying representation real line such that the ordering is different. With assumption $A_{(n)}$ up to $A_{(n+T-1)}$ and all observations being interchangeable, $\binom{n+T}{n}$ way of distribution are assumed to be equally likely.

So given data (n, j), the lower probability $\underline{p}_{(n,j)}$ for observing r occurrences of "B" within any future T stages count all the ways of distribution such that there must be r units out of T units of future observation Y_w that are less than ℓ and the upper probability $\overline{p}_{(n,j)}$ for observing r occurrences of "B" within any future T stages count all the ways of distribution such that there could be r units out of T units of future observation Y_w that are less than ℓ. Hence:

$$\overline{p}_{(n,j)}(S_T \in \{r\}) = \binom{n+T}{n}^{-1} \times \left[\binom{j+r}{r}\binom{n-j+T-r}{T-r}\right] \tag{2.2.18}$$

$$\underline{p}_{(n,j)}(S_T \in \{r\}) = \binom{n+T}{n}^{-1} \times \left[\binom{j-1+r}{r}\binom{n-j-1+T-r}{T-r}\right] \tag{2.2.19}$$

With same counting argument, Coolen [8, 13, 18] also gives the formulas of NPI imprecise probabilities for other form of future random quantity S_T which is summaries below:

The most general form of S_T is $S_T \in \{z_i\}_{i=1}^{i=\alpha}$, $\alpha \leq T$ with $0 \leq z_i < z_j \leq T$ for $i < j$

$$\overline{p}_{(n,j)}(S_T \in \{z_i\}_{i=1}^{i=\alpha}) = \binom{n+T}{n}^{-1} \times \sum_{i=1}^{\alpha}\left[\binom{j+z_i}{z_i} - \binom{j+z_{i-1}}{z_{i-1}}\right]\binom{n-j+T-z_i}{T-z_i} \tag{2.2.20}$$

$$\underline{p}_{(n,j)}(S_T \in \{z_i\}_{i=1}^{i=\alpha}) = 1 - \overline{p}_{(n,j)}(S_T \in \mathbb{N}_0^T \setminus \{z_i\}_{i=1}^{i=\alpha}) \tag{2.2.21}$$

Let $\mathbb{N}_{i_1}^{i_2}$ denote the set of natural number from i_1 to i_2 where $i_1 < i_2$ and $i_1, i_2 \in \mathbb{N}$.

Then for $S_T \in \mathbb{N}_0^m$ and $S_T \in \mathbb{N}_m^T$,

$$\overline{p}_{(n,j)}(S_T \in \mathbb{N}_m^T) = \binom{n+T}{n}^{-1} \times \sum_{i=m}^{T} \left[\binom{j+i}{i}\binom{n-j-1+T-i}{T-i}\right] \tag{2.2.22}$$

$$\underline{p}_{(n,j)}(S_T \in \mathbb{N}_m^T) = \binom{n+T}{n}^{-1} \times \sum_{i=m}^{T} \left[\binom{j-1+i}{i}\binom{n-j+T-i}{T-i}\right] \tag{2.2.23}$$

$$\overline{p}_{(n,j)}(S_T \in \mathbb{N}_0^m) = \binom{n+T}{n}^{-1} \times \sum_{i=0}^{m} \left[\binom{j-1+i}{i}\binom{n-j+T-i}{T-i}\right] \tag{2.2.24}$$

$$\underline{p}_{(n,j)}(S_T \in \mathbb{N}_0^m) = \binom{n+T}{n}^{-1} \times \sum_{i=0}^{m} \left[\binom{j+i}{i}\binom{n-j-1+T-i}{T-i}\right] \tag{2.2.25}$$

And for $S_T \in \mathbb{N}_{m_1}^{m_2}$,

$$\overline{p}_{(n,j)}(S_T \in \mathbb{N}_{m_1}^{m_2}) = \overline{p}_{(n,j)}(S_T \in \mathbb{N}_{m_1}^T) - \underline{p}_{(n,j)}(S_T \in \mathbb{N}_{m_2+1}^T) \tag{2.2.26}$$

$$= \overline{p}_{(n,j)}(S_T \in \mathbb{N}_0^{m_2}) - \underline{p}_{(n,j)}(S_T \in \mathbb{N}_0^{m_1-1}) \tag{2.2.27}$$

$$\underline{p}_{(n,j)}(S_T \in \mathbb{N}_{m_1}^{m_2}) = \binom{n+T}{n}^{-1} \times \sum_{i_1=m_1}^{m_2} \sum_{i_2=i_1}^{m_2} \left[\binom{j-1+i_1}{i_1}\binom{n-j-1+T-i_2}{T-i_2}\right] \tag{2.2.28}$$

The current imprecise probability formulas in NPI for Bernoulli data allow one to compute the lower and upper expectation of monotonic function of future random quantity S_T. This is achieved by constructing the lower expectation measure $p_{\underline{E}(f)}$ and the upper expectation measure $p_{\overline{E}(f)}$ via following formulas.

Denote monotonically increasing function as $f_{\uparrow}(\cdot)$ and monotonically decreasing function as $f_{\downarrow}(\cdot)$

For $f_{\uparrow}(\cdot)$, to find $p_{\underline{E}(f)}(\cdot)$, one assigns the least possible mass to the greatest possible value of S_T, thus,

$$\begin{aligned} p_{\underline{E}(f)}(f_{\uparrow}(S_T = m)) &= \underline{p}_{(n,j)}(S_T \geq m) - \underline{p}_{(n,j)}(S_T \geq m+1) \qquad (2.2.29) \\ &= \overline{p}_{(n,j)}(S_T \leq m) - \overline{p}_{(n,j)}(S_T \leq m-1) \quad \forall m \in \mathbb{N}_0^T \\ &= \binom{n+T}{n}^{-1} \times \left[\binom{j-1+m}{m}\binom{n-j+T-m}{T-m}\right] \end{aligned}$$

And to find $p_{\overline{E}(f)}(\cdot)$, one assigns the greatest possible mass to the greatest possible value of S_T, thus,

$$\begin{aligned} p_{\overline{E}(f)}(f_\uparrow(S_T = m)) &= \underline{p}_{(n,j)}(S_T \leq m) - \underline{p}_{(n,j)}(S_T \leq m-1) \\ &= 1 - \overline{p}_{(n,j)}(S_T \geq m+1) - 1 + \overline{p}_{(n,j)}(S_T \geq m) \\ &= \overline{p}_{(n,j)}(S_T \geq m) - \overline{p}_{(n,j)}(S_T \geq m+1) \quad \forall m \in \mathbb{N}_0^T \\ &= \binom{n+T}{n}^{-1} \times \left[\binom{j+i}{m} \binom{n-j-1+T-m}{T-m} \right] \end{aligned} \tag{2.2.30}$$

For $f_\downarrow(\cdot)$, to find $p_{\underline{E}(f)}(\cdot)$, one assigns the greatest possible mass to the greatest possible value of S_T, thus,

$$\begin{aligned} p_{\underline{E}(f)}(f_\downarrow(S_T = m)) &= \underline{p}_{(n,j)}(S_T \leq m) - \underline{p}_{(n,j)}(S_T \leq m-1) \\ &= 1 - \overline{p}_{(n,j)}(S_T \geq m+1) - 1 + \overline{p}_{(n,j)}(S_T \geq m) \\ &= \overline{p}_{(n,j)}(S_T \geq m) - \overline{p}_{(n,j)}(S_T \geq m+1) \quad \forall m \in \mathbb{N}_0^T \\ &= \binom{n+T}{n}^{-1} \times \left[\binom{j+i}{m} \binom{n-j-1+T-m}{T-m} \right] \end{aligned} \tag{2.2.31}$$

And to find $p_{\overline{E}(f)}(\cdot)$, one assigns the least possible mass to the greatest possible value of S_T, thus,

$$\begin{aligned} p_{\overline{E}(f)}(f_\downarrow(S_T = m)) &= \underline{p}_{(n,j)}(S_T \geq m) - \underline{p}_{(n,j)}(S_T \geq m+1) \\ &= \overline{p}_{(n,j)}(S_T \leq m) - \overline{p}_{(n,j)}(S_T \leq m-1) \quad \forall m \in \mathbb{N}_0^T \\ &= \binom{n+T}{n}^{-1} \times \left[\binom{j-1+m}{m} \binom{n-j+T-m}{T-m} \right] \end{aligned} \tag{2.2.32}$$

However, the current existing imprecise probability formulas in NPI for Bernoulli data are unable to compute imprecise expectation for a general function of the future random quantity S_T. Moreover, NPI for Bernoulli data so far is only able to handle precise Bernoulli data. These two unsolved challenging issues give part of the motivations of this thesis and will be addressed in Chapter 3.

2.3 Financial objects, concepts and terminologies

In this section, with reasonable assumptions, under the binomial tree, relevant financial objects are defined which be used in later NPI application. Related financial concepts are also introduced. In the end, some financial terminologies are explained which provides necessary information for one who is less familiar in finance.

2.3.1 Financial objects and concepts

Since the attempt in this thesis is to applied NPI for Bernoulli data in finance, all the financial objects of interest are defined under the binomial tree with reasonable assumptions.

[Asset]: Through the thesis, an asset normally refers to a stock of which has sufficient participant in trading. We are only interested in the asset price at time T. The asset price at time T is treated as a random variable $A_T(S_T)$ which depends on another random variable S_T. S_T is a sum of T units of Bernoulli random variable $X_t \in \{0, 1\}$ which indicates whether the asset price goes up at time t. Thus,

$$S_T = \sum_{t=1}^{T} X_t$$

The relation between $A_T(S_T)$ and S_T is stated as followed:

$$A_T(S_T) = a_0 u^{S_T} d^{T-S_T}$$

where a_0 is a fixed value, representing the initial asset price at time $T = 0$; u is the magnitude of upward movement in each time stage, $u \in (1, +\infty)$; d the magnitude of down movement in each time stage, $d \in (0, 1)$; Figure 2.3 provide a graphical illustration.

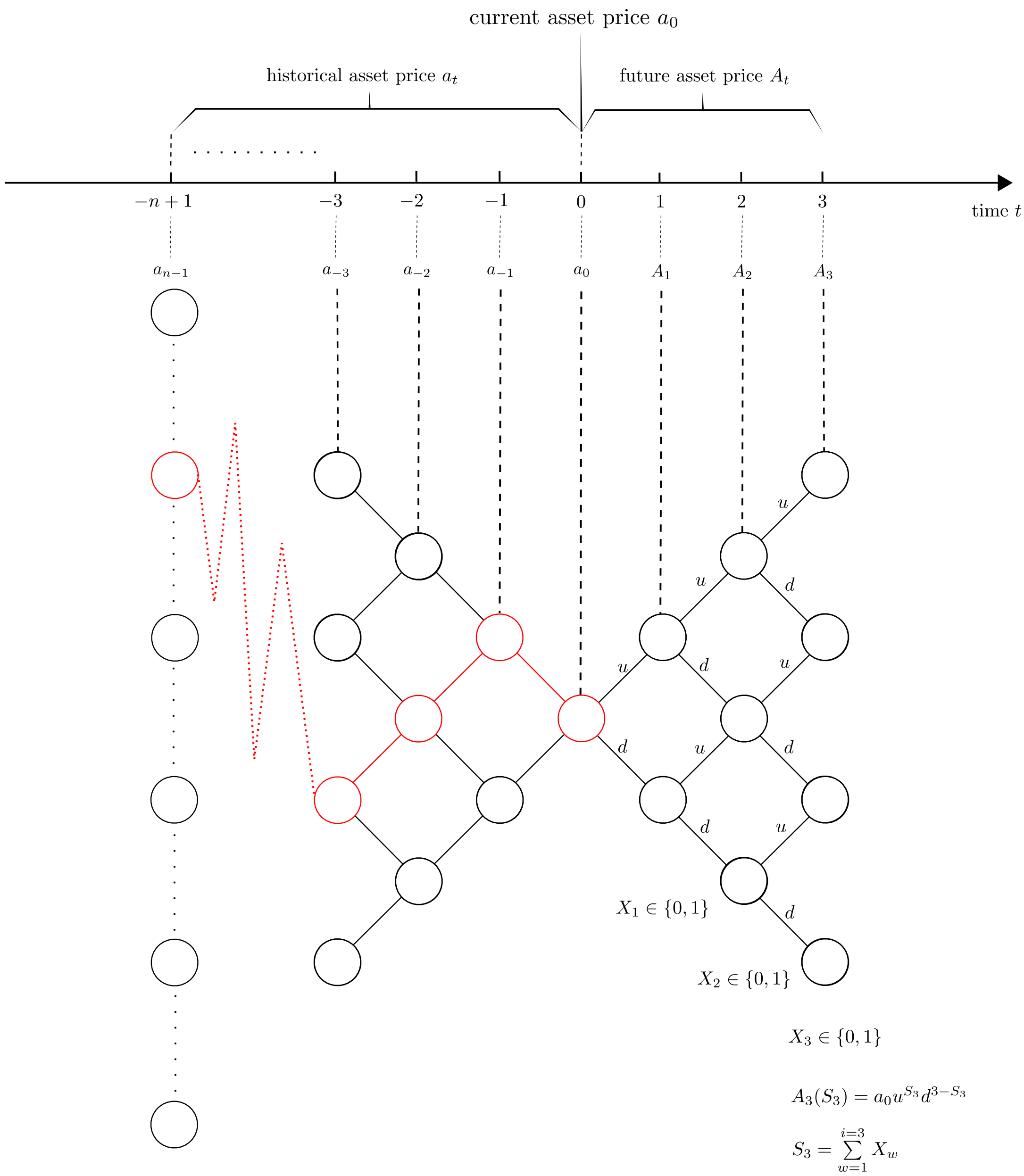

Figure 2.3: Asset price A_t at time $t = 3$ in the binomial tree.

Some reasonable assumptions about u, d, and random variable X_t are made, which renders NPI for Bernoulli data a suitable statistical methodology for later financial applications.

In a short period of time interval, the trading participants upon a certain asset are assumed to be relatively same and stable. On aggregate, trading behaviors within a short period of time could be assumed to be probabilistically homogeneous. Based on this:

1) u and d are assumed to be constant.

2) The probability distribution p_t of random variable X_t within a short period of time is assumed to be "stable". In other words, within a short period of time interval (a, b), it is assumed that:

$$\sup_{t\in(a,b)} p_t(X_t \in w) - \inf_{t\in(a,b)} p_t(X_t \in w) < \epsilon \quad \forall w \subset \{0, 1\} \quad \text{for small} \ \ \epsilon$$

Therefore, within a short period of time, the sequence of X_t is approximately exchangeable. Consequently, NPI for Bernoulli data is suitable method make inference about X_t and S_T, thus the asset price A_T at time T.

[Risk free interest rate r]: Throughout the thesis, it is assumed there exists a constant continuously compounding risk free interest rate r in the market for one to invest or borrow cash. In order to make the asset an indeterministic choice to buy or short sell, the relation $0 < d < e^r < u$ is required

[Present value & discount factor]: The existence of risk free interest rate allows one to compare monetary values which is at different time stage. This is achieved by calculating the present value PV of the monetary value f_t at time t via discount factor $B(t) = e^{rt}$, which is:

$$PV(f_t) = f_t \times B(t)^{-1} = f_t e^{-rt}$$

[Financial portfolio]: A financial portfolio is a set of financial objects that one own or owe.

[Financial derivative of an asset] A financial derivative of an asset is a contract between two parties of which the payoff/loss at a specific time depends on the asset-property over a time interval. Mathematically speaking, let T be the end time of a financial derivative. The payoff or loss of a financial derivative of an asset A_t at time T is a function $f(A_t, t \in \{0, T\})$ whose value depend on the asset price A_t over the time interval $t \in \{0, T\}$. The financial derivative of interest in this thesis are introduced below.

[European option of a asset] Let x^+ denote the maximum value between x and 0, namely

$$x^+ = max(x, 0)$$

A European call option gives the holder the right but not obligation to buy the underlying asset on a certain date T for a certain price K [32, 33]. There are some equivalent terminologies for "certain date" and "certain price" in the financial industry. "Expiration date" and "exercise date", "maturity date" are frequently the equivalent terminologies for "certain date" while "exercise price" and "strike price" are the equivalent terminologies for "certain price".

Mathematically, the value of European call option with maturity date T and strike price K is a function of the underlying asset price A_t. $\Lambda_c(A_t, K)$ with boundary condition at time T as

$$\Lambda_c(A_T, K) = (A_T - K)^+$$

For an underlying asset price A_t evolving as the above description in the binomial tree, Cox, Ross, & Rubinstein (CRR) developed the binomial options pricing model in 1979 [25]. In the CRR model, by replicating the performance of the European option with a self-financing portfolio in each time step, one can find that there exists a unique arbitrage-free price $\Lambda_c^{\mathbb{Q}}(A_t, K)$ of the call option at time t ($t < T$) with above boundary condition. And the unique arbitrage-free price $\Lambda_c^{\mathbb{Q}}(A_t, K)$ of the call option at time t can be computed via a risk neutral measure $\mathbb{Q}$.

In $\mathbb{Q}$ measure, the "risk free" probability of "going up" in each time stage is $q = \frac{e^r - d}{u - d}$ for each time stage. For a call option with above boundary condition, one has:

$$\begin{aligned}
\Lambda_c^{\mathbb{Q}}(A_t(S_t), K) &= B(T-t)^{-1} E_{\mathbb{Q}}(\Lambda_c(A_T(S_T), K)|\mathscr{A}_t) \\
&= B(T-t)^{-1} \sum_{i=0}^{T-t} \binom{T-t}{i} (A_t u^i d^{T-t-i} - K)^+ q^i (1-q)^{T-t-i}
\end{aligned}$$

A European put option gives holder the right but not obligation to sell the underlying asset on a certain date for a certain price. [32, 33]

Mathematically, the value of European put option with maturity date T and strike price K is a function of underlying asset price A_t. $\Lambda_p(A_t, t)$ with boundary condition at time T as

$$\Lambda_p(A_T, K) = (K - A_T)^+$$

With the similar replication argument, CRR model also showed that there existed a unique arbitrage-free price $\Lambda_p^{\mathbb{Q}}(A_t, K)$ for the put option at time t with above boundary condition. And $\Lambda_p^{\mathbb{Q}}(A_t, K)$ could be computed via the same $\mathbb{Q}$ measure as above:

$$\begin{aligned}
\Lambda_p^{\mathbb{Q}}(A_t(S_t), K) &= B(T-t)^{-1} E_{\mathbb{Q}}(\Lambda_p(A_T(S_T), K)|\mathscr{A}_t) \\
&= B(T-t)^{-1} \sum_{i=0}^{T-t} \binom{T-t}{i} (K - A_t u^i d^{T-t-i})^+ q^i (1-q)^{T-t-i}
\end{aligned}$$

In the CRR model, there is no real probability or "risk involved" in the derivation of arbitrage-free price. The $\mathbb{Q}$ measure is not real probability but a convenient way to compute arbitrage-free price. Although in a complete market, anyone who is willing to buy at price $y_t > \Lambda^{\mathbb{Q}}(A_t, K)$ or sell at price $y_t < \Lambda^{\mathbb{Q}}(A_t, K)$ for the European option with boundary condition $\Lambda(A_T, K)$ at time t will become a free money source for an arbitrager, this behaviour is still rational if:

1) At time t one's personally expected present value of the European option payoff under one's risk measure $\mathbb{P}$ is greater than the arbitrage-free price of the option, $\Lambda^{\mathbb{P}}(A_t, K) = B_{T-t}^{-1} E_{\mathbb{P}}(\Lambda(A_T, K)|\mathscr{A}_t) > \Lambda^{\mathbb{Q}}(A_t, K)$, or under one's risk measure, the probability of the event $\Lambda(A_T, K) > \Lambda^{\mathbb{Q}}(A_t, K) B_{T-t}$ at time t is greater than a thresold value, when one considers to buy.

2)One's expected present value of the European option under one's risk measure is less than the arbitrage-free price of the option, $\Lambda^{\mathbb{P}} = B_{T-t}^{-1} E_{\mathbb{P}}(\Lambda(A_T, K)|\mathscr{A}_t) < \Lambda^{\mathbb{Q}}(A_t, K)$, or under one's risk measure, the probability of the event $\Lambda(A_T, K) < \Lambda^{\mathbb{Q}}(A_t, K) B_{T-t}$ at time t is greater than a thresold value when one considers to sell.

In both situations, under one's risk measure, one is confident enough to make positive payoff at time T either expectationally or probabilistically.

2.3.2 Financial terminologies

[Short sell/selling]: Short selling is a financial action when one sells a financial object which has monetary value but is not owned by the person. One will receive the cash equal to the price of the financial object at the time one short sells and one is obligated to buy back the financial object and return to the owner upon a specific time. Short selling is a rational behaviour when one anticipate the price of the object is likely to going down in the future time. By short selling, one can exploit the potential profit from the price decrease of the financial object in the future time.
[Enter a (risk) position]: When one buys or short sells a financial object, we say one enters a "risk position", or simply enter a position of this financial object. When one enters a "position" of an object, one's capital gain or loss becomes random and it will depend on the contingent price change of the object until the time one closes the position.
[Close a (risk) position]: After one enters a position by the mean of buying or shorting selling, one executes the reverse action by selling the object or buying the object returning to owner, we said one close the "risk position" or simply the "position" of the object. Risk positions can be further categorised into short or long position as below.
[Short position]: When one short sells a financial object which it is not owned by oneself, we say one enters the "short position" of this object. One is then obligated to buy back the object and return to the owner upon a specific time.
[Long position]: When one buys a financial object, we say one enters the long position. One is then anticipating that the price of the object would increase in the future.

Mathematically, given initial capital C, when one keeps all one's cash risk free investment, one's capital gain $\Delta(T) = C(e^{rT} - 1)$ is a deterministic function of time T.

In any time stage, when one enters multiple positions of different financial object using part of one's total cash, then one's capital gain or loss at any future time is a linear combination of multiple stochastic processes and the deterministic function of the remaining cash one invest in risk free rate.

Chapter 3

Further development of NPI for Bernoulli data

In this chapter, based on the imprecise probability definitions in Chapter 2, a general algorithm to construct imprecise expectation measures for a general function of a finite random variable is presented. Subsequently, using the underlying latent variable representation of NPI, the mass function of NPI for Bernoulli data is constructed. With a one to one mapping between the mass function value and path counting within the lattice, the constructed mass function is verified that it satisfies the mass function defined in Definition 2.1.3 and produce same imprecise probability value as mentioned in the previous chapter. This constructed mass function also leads to a new formula in lower probability and its consistence is proved. With the proposed general algorithm and the constructed mass function, one is now able to find the imprecise expectation measure for a general function of multiple future stages observations S_T. An example of how to apply the algorithm is presented.

3.1 Greedy mass assignment algorithm

A greedy mass assignment algorithm (GMA) is presented in this section for one to construct imprecise expectation measures for a general function of a finite random variable on an imprecise probability space.

Consider an imprecise probability space $[\Omega, \mathscr{A} = \sigma(L^{-1}), m(\cdot)]$, where L is a finite random variable $L : \Omega \longrightarrow F$, F is a finite set with N elements and a function $f : F \longrightarrow M$, M is a finite order set, therefore f could induces a order $\succ$ on F by the order of M.

With the induced order $\succ$ on F, F now can be written as $F = \{i_\alpha\}_{\alpha=1}^{\alpha=N}$, and $i_k \succ i_j$ for $k > j$.

Let $\mathcal{P}(\cdot)$ denote the power set operator and let $Q_y = (Q_y^1, Q_y^2, Q_y^3, Q_y^4)$ denote the yth stage of the algorithm, where Q_y^1 denotes the residual elements in F that have not been yet assigned mass value; Q_y^2 denotes the current highest order element in Q_y^1, if there is a tie, then they are equally the highest order elements; Q_y^3 denotes residual sets in $\mathcal{P}(F)$ of which the mass value have not yet been used; Q_y^4 denotes records of elements in F that have been assigned a probability mass.

The algorithm moves from yth stage to $(y+1)$-th stage in following way:

0. Check if $Q_y^1 = \emptyset$, if $Q_y^1 = \emptyset$, the algorithm stops, else if $Q_y^1 \neq \emptyset$ proceed with following steps.

1. Use the mass function to evaluate and assigned the current maximum possible mass value $\sum\limits_{\substack{\epsilon \in Q_y^3 \\ \epsilon \cap Q_y^2 \neq \emptyset}} m(\epsilon)$ to current highest order element Q_y^2.

2. Record the mass assignment in Q_{y+1}^4, namely $Q_{y+1}^4 = Q_y^4 \cup \{(Q_y^2, \sum\limits_{\substack{\epsilon \in Q_y^3 \\ \epsilon \cap Q_y^2 \neq \emptyset}} m(\epsilon))\}$.

3. Move to next stage $Q_{y+1} = (Q_{y+1}^1, Q_{y+1}^2, Q_{y+1}^3, Q_{y+1}^4)$, where $Q_{y+1}^1 = Q_y^1 \setminus Q_y^2$, $Q_{y+1}^2 =$ highest order element in Q_{y+1}^1 with respect to the defined order $\succ$, $Q_{y+1}^3 = Q_y^3 \setminus \{\epsilon | \epsilon \in Q_y^3, \epsilon \cap Q_y^2 \neq \emptyset\}$ and $Q_{(y+1)}^4 = Q_y^4 \cup \{(Q_y^2, \sum\limits_{\substack{\epsilon \in Q_y^3 \\ \epsilon \cap Q_y^2 \neq \emptyset}} m(\epsilon))\}$.

Assuming the most general case that the mass function has non zero mass value for all elements in $\mathcal{P}(F)$ and there is no tie in the F with respective to the order $\succ$, the algorithm initiates at stage $Q_0 = (Q_0^1, Q_0^2, Q_0^3, Q_0^4)$ with $Q_0^1 = F = \{i_\alpha\}_{\alpha=1}^{\alpha=N}$, $Q_0^2 = i_N$, $Q_0^3 = \mathcal{P}(F)$, $Q_0^4 = \emptyset$. After N iterations, the algorithm will stop and one thus find the upper expectation measure from Q_N^4.

To find the lower expectation measure, one simply need to use the reversed order of $\succ$ in the algorithm.

3.2 Mass function of NPI for Bernoulli data

In this section, the goal is to construct mass function in NPI for Bernoulli data. To achieve it, the sample space and event space of NPI for Bernoulli data is firstly specified below:

Let $A^{<\infty}$ denotes finite Cartesian product of the set A.

The sample space Ω in NPI for Bernoulli data is then $\Omega = \{B, B^c\}^{<\infty} = \{B, B^c\}_1 \times \{B, B^c\}_2 \times ... \times \{B, B^c\}_t \times ... \times \{B, B^c\}_{n<\infty}$.

The event space $\mathscr{A}$ of NPI for Bernoulli data for a single future stage observation is $\mathscr{A} = \sigma(X_t^{-1})$, where $\sigma(\cdot)$ denotes the operator which generates the smallest sigma algebra using a collection of set inside the argument $(\cdot)$ and X_t is a Bernoulli random variable on future t-th stage. $X_t : \Omega \longrightarrow \{0, 1\}$, $O \in \Omega$

$$X_t(O) = \mathbb{1}_{\{B\}_t}(O)$$

The event space $\mathscr{A}$ of NPI for Bernoulli data for multiple future stages observations is $\mathscr{A} = \sigma(S_T^{-1})$, where S_T is a sum of T units of Bernoulli random variables X_t on different future stages.

$$S_T = \sum_{t \in W} X_t \quad \text{where} \quad W \subset \mathbb{N}^+ \quad \text{and} \quad |W| = T$$

3.2.1 NPI for Bernoulli data in a single future stage observation

Assume $A_{(n)}$, given data (n, j) in past n history observation, the mass function $m_{(n,j)}$ of NPI for Bernoulli data for space $[\Omega = \{B, B^c\}^{<\infty}, \mathscr{A} = \sigma(X_t^{-1})]$. $\forall w \in \mathbb{N}$ could be constructed as:

$$m_{(n,j)}(X_t \in \{1\}) = m(Y_t < \ell) = \tfrac{j}{n+1}$$
$$m_{(n,j)}(X_t \in \{0\}) = m(Y_t > \ell) = \tfrac{n-j}{n+1}$$
$$m_{(n,j)}(X_t \in \{0, 1\}) = m(Y_t < \ell \ \text{ or } \ Y_t > \ell) = \tfrac{1}{n+1}$$
$$m_{(n,j)}(X_t \in \emptyset) = m(Y_t \in \emptyset) = 0$$

where Y_t is the latent variable representation of X_t, and the mass function value for $X_t \in \{0,1\}$ is given by the $A_{(n)}$ assumption in the situation where the order of latent variable Y_t and threshold latent variable ℓ is unknown. (See Figure 3.1.)

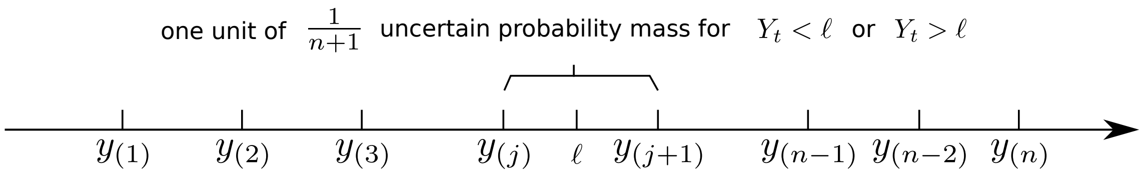

Figure 3.1: NPI for Bernoulli data in single future stage observation

It could be verified that the constructed mass function above satisfies the Definition 2.1.3. ($m_{(n,j)}(\emptyset) = 0$; $m_{(n,j)}(\epsilon) \geq 0$,$\forall \epsilon \in \sigma(X_t^{-1})$; $\sum_{\epsilon \in \mathscr{A}} m_{(n,j)}(\epsilon) = 1$)

Moreover, with imprecise probability $\underline{p}$ and $\overline{p}$ defined in Definition 2.1.3, the constructed mass function above yields identical imprecise probability values for X_i as shown in previous chapter Equality 2.2.14-2.2.17

For example, given data (n, j), the imprecise probability of the event that future t-th stage observation is "B", or equivalently "$X_t \in \{1\}$" is :

$$
\begin{aligned}
&\overline{\underline{p}}_{(n,j)}(X_t \in \{1\}) \\
&= \left[\underline{p}_{(n,j)}(X_t \in \{1\}), \overline{p}_{(n,j)}(X_t \in \{1\})\right] \\
&= \left[\sum_{\substack{\epsilon \in \mathscr{A} \\ \epsilon \subset \{B\}_t}} m_{(n,j)}(\epsilon), \sum_{\substack{\epsilon \in \mathscr{A} \\ \epsilon \cap \{B\}_t \neq \emptyset}} m_{(n,j)}(\epsilon)\right] \\
&= \left[m_{(n,j)}(X_t \in \{1\}) + m_{(n,j)}(X_t \in \emptyset)\ ,\ m_{(n,j)}(X_t \in \{1\}) + m_{(n,j)}(X_t \in \{0,1\})\right] \\
&= \left[\frac{j}{n+1}, \frac{j+1}{n+1}\right]
\end{aligned}
$$

Similarly, one can have

$$
\begin{aligned}
\overline{\underline{p}}_{(n,j)}(X_t \in \{0\}) &= \left[\frac{n-j}{n+1}, \frac{n-j+1}{n+1}\right] \\
\overline{\underline{p}}_{(n,j)}(X_t \in \{0,1\}) &= [1, 1] \\
\overline{\underline{p}}_{(n,j)}(X_t \in \emptyset) &= [0, 0]
\end{aligned}
$$

3.2.2 NPI for Bernoulli data in multiple future stages observations

Assume $A_{(n)}$ up to $A_{(n+T-1)}$, given data (n, j), the mass function $m_{(n,j)}$ of NPI for Bernoulli data NPI for the space $[\Omega = \{B, B^c\}^{<\infty}, \mathscr{A}_T = \sigma(S_T^{-1})]$ could be constructed as followed:

Define operator $\mathscr{C}$ which generates a collection of subset which contains only one value or several consecutive positive values in a consecutive positive integer set. Then,

$$\mathscr{C}(\mathbb{N}_{i_1}^{i_2}) = \{\mathbb{N}_{j_1}^{j_2} | j_1, j_2 \in \mathbb{N}_{i_1}^{i_2}, i_1 \le j_1 \le j_2 \le i_2\}.$$

Denote the set $\mathscr{B}_T = \{S_T^{-1}(a) | a \in \mathscr{C}(\mathbb{N}_0^T)\}$.

The NPI mass function $m_{(n,j)}(\cdot) : \mathscr{A}_T \to [0, 1]$ only assigns non-zero value to element in the set $\mathscr{B}_T = \{S_T^{-1}(a) | a \in \mathscr{C}(\mathbb{N}_0^T)\}$. The rest of elements in $\mathscr{A}_T$ have zero mass.

$$m_{(n,j)}(S_T \in \epsilon) = \begin{cases} \binom{j-1+r_1}{r_1}\binom{n-j-1+T-r_2}{T-r_2} \times \binom{n+T}{n}^{-1} & \text{for } \epsilon = \mathbb{N}_{r_1}^{r_2} \in \mathscr{C}(\mathbb{N}_0^T), \\ 0 & \text{otherwise.} \end{cases} \tag{3.2.1}$$

The mass value for $S_T \in \mathbb{N}_{r_1}^{r_2}$ can be interpreted as the shared mass that r_1 up to r_2 out of T future observations are successes. In other words, at least r_1 out of T future observations are successes, and at least $T - r_2$ out of T future observations are failures, and the rest $r_2 - r_1$ out of T future observations are uncertain. So the corresponding way of distribution in the latent representation is as Figure 3.2.

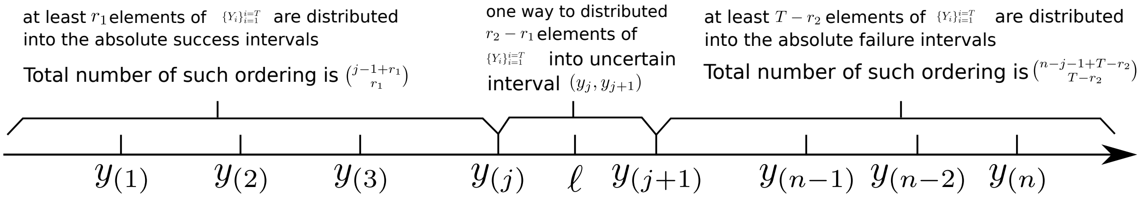

Figure 3.2: NPI for Bernoulli data in multiple future stages observations

Ahmad [1] found that there is a one to one mapping between NPI imprecise probability for S_T and lattice path counting in the $n \times T$ lattice. The above mass function construction also has its corresponding lattice path counting in the $n \times T$ lattice. We used this fact to verify that the constructed mass function $m_{(n,j)}$ satisfied Definition 2.1.3 and also produces the same imprecise probability values for S_T as mentioned in the previous chapter.

In a $n \times T$ lattice, the total number of paths from (0,0) to (n,T) which allow only upward and rightward movement is $\binom{n+T}{n}$.

With data (n, j), the mass value for $S_T \in \mathbb{N}_{r_1}^{r_2}$ is $\binom{j-1+r_1}{r_1}\binom{n-j-1+T-r_2}{T-r_2}$, which corresponds to number of paths from $(0, 0)$ to (n, T) that must pass through $(j - 1, r_1)$, (j, r_1), (j, r_2),$(j + 1, r_2)$, indicated as the red tunnel in the Figure 3.3.

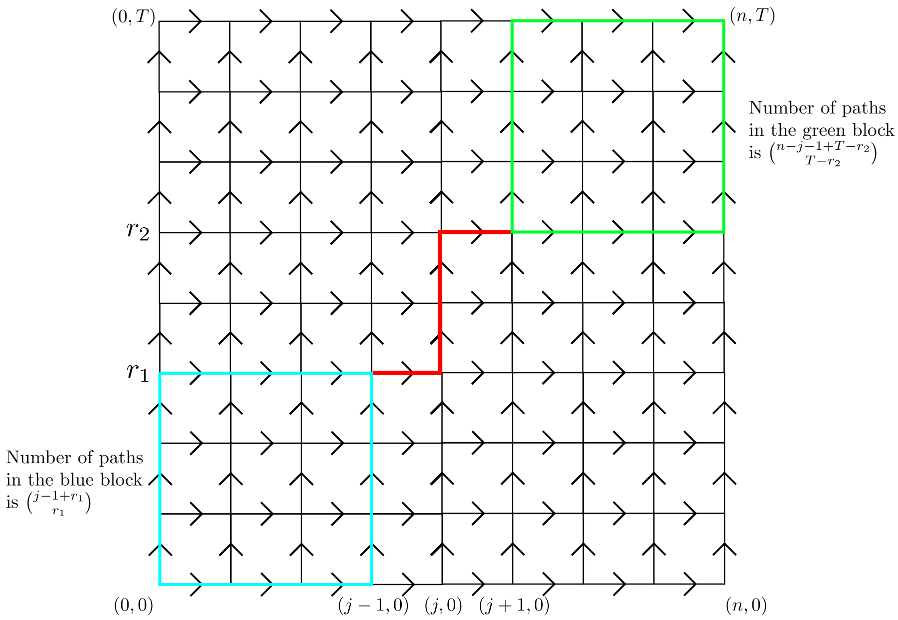

Figure 3.3: $\binom{j-1+r_1}{r_1}\binom{n-j-1+T-r_2}{T-r_2}$ paths from (0,0) to (0,T) passing the red tunnel which corresponds to the mass value for $\mathbb{N}_{r_1}^{r_2}$

Now, let's verify the above mass function $m_{(n,j)}$ construction satisfied Definition 2.1.3

First, $\emptyset \notin \mathscr{B}_T$, thus $m_{(n,j)}(\emptyset) = 0$ and $\forall \epsilon \in \sigma(S_T^{-1})$, we always have non negative number of path counting, thus $m_{(n,j)}(\epsilon) \geq 0$

Second, one need to show $\sum\limits_{\epsilon \in \mathscr{A}_T} m_{(n,j)}(\epsilon) = 1$

$$\sum_{\epsilon\in\mathscr{A}_T} m_{(n,j)}(\epsilon) = \sum_{\epsilon\in\mathscr{B}_T} m_{(n,j)}(\epsilon) = \sum_{\epsilon\in\mathscr{C}(\mathbb{N}_0^T)} m_{(n,j)}(S_T \in \epsilon = \mathbb{N}_{r_2}^{r_1})$$

$$= \sum_{\substack{r_1,r_2\in\mathbb{N}_0^T \\ 0\leq r_1\leq r_2\leq T}} \binom{j-1+r_1}{r_1}\binom{n-j-1+T-r_2}{T-r_2} \times \binom{n+T}{n}^{-1}$$

$$= \underbrace{\sum_{r_1=0}^{T}\sum_{r_2=r_1}^{T} \binom{j-1+r_1}{r_1}\binom{n-j-1+T-r_2}{T-r_2}}_{} \times \binom{n+T}{n}^{-1}$$

"This quantity is number of paths that have to pass through $(r_1, j-1)$ and $(r_2, j+1)$ for some $0 \leq r_1 \leq r_2 \leq T$ which is number of all the paths from $(0,0)$ to (n,T)"

$$= \binom{n+T}{n} \times \binom{n+T}{n}^{-1}$$

$$= 1$$

To verify the constructed mass function $m_{(n,j)}$ produces same imprecise probability value for S_T, we used the fact from Ahmad's paper [1] that the lower probability for $S_T \in \mathbb{N}_{r_1}^{r_2}$ is total number of paths from $(0,0)$ to (n,T) that enter in any of $\{(j-1,i)\}_{i=r_1}^{i=r_2}$ channels and come out from $\{(j+1,i)\}_{i=r_1}^{i=r_2}$ channels. (See Figure 3.4 indicated by red colour.)

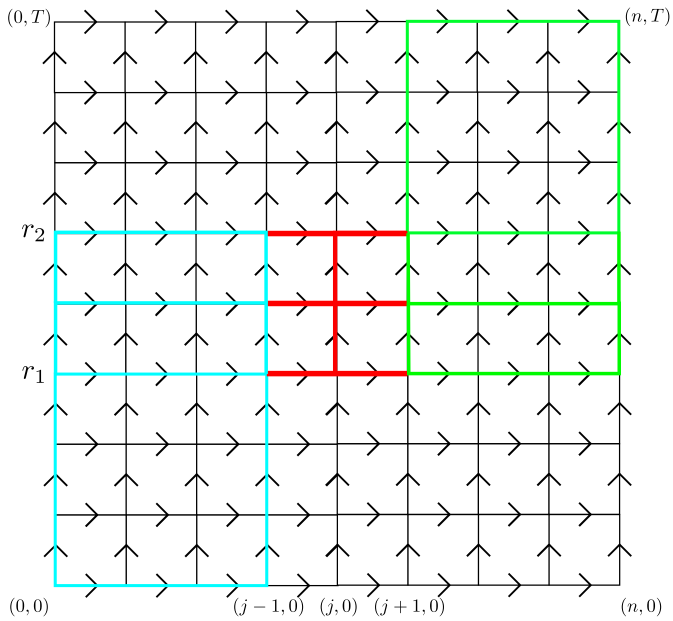

Figure 3.4: Paths that go through red channels correspond to lower probability mass assignment

So the lower probability $\underline{p}_{(n,j)}$ for $S_T \in \mathbb{N}_{r_1}^{r_2}$ is

$$\underline{p}_{(n,j)}(S_T \in \mathbb{N}_{r_1}^{r_2}) \quad = \quad \binom{n+T}{n}^{-1} \times \sum_{i_1=r_1}^{r_2} \sum_{i_2=i_1}^{r_2} \left[\binom{j-1+i_1}{i_1} \binom{n-j-1+T-i_2}{T-i_2} \right]$$

By the definition of lower probability in the previous chapter, the constructed mass function also yield the same lower probability value for $S_T \in \mathbb{N}_{r_1}^{r_2}$ as they have the same pathing counting principle in the lattice.

$$\sum_{\substack{\epsilon \in \mathscr{C}(\mathbb{N}_0^T) \\ \epsilon \subset \mathbb{N}_{r_1}^{r_2}}} m_{(n,j)}(S_T \in \epsilon) = \binom{n+T}{n}^{-1} \times \sum_{i_1=r_1}^{r_2} \sum_{i_2=i_1}^{r_2} \left[\binom{j-1+i_1}{i_1} \binom{n-j-1+T-i_2}{T-i_2} \right]$$

Similarly for the upper probability definition, consider $S_T \in \mathbb{N}_{r_1}^{r_2}$

$$\begin{aligned}
&\overline{p}_{(n,j)}(S_T \in \mathbb{N}_{r_1}^{r_2}) \\
=& \binom{n+T}{n}^{-1} \times \left[\sum_{i=r_1}^{T} \binom{j+i}{i} \binom{n-j-1+T-i}{T-i} - \sum_{i=r_2+1}^{T} \binom{j-1+i}{i} \binom{n-j+T-i}{T-i} \right] \\
=& \sum_{\substack{\epsilon \in \mathscr{C}(\mathbb{N}_0^T) \\ \epsilon \cap \mathbb{N}_{r_1}^{r_2} \neq \emptyset}} m_{(n,j)}(S_T \in \epsilon)
\end{aligned}$$

Both values are equal to the total number of paths from $(0,0)$ to (n,T) that enter in any of $\{(j-1,i)\}_{i=0}^{i=r_2}$ channels and come out from any of $\{(j+1,i)\}_{i=r_1}^{i=T}$ channels. (See Figure 3.5 indicated by red colour.)

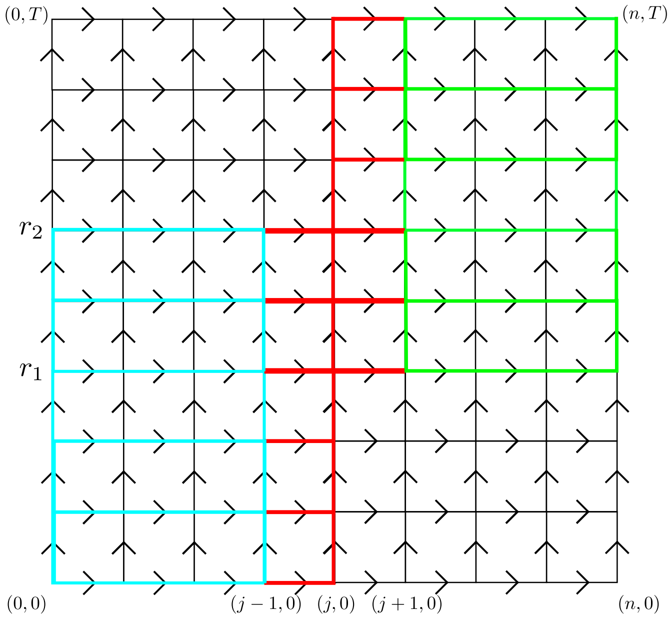

Figure 3.5: Paths that go through red channels correspond to the upper probability mass assignment

With the constructed mass function, one now can directly compute the lower probability for the most general form of S_T instead of using the conjugacy property as Formula 2.2.21 in previous chapter.

Recall the most general form of S_T is $S_T \in \{z_i\}_{i=1}^{i=\alpha}$, $\alpha \leq T$ with $0 \leq z_i < z_j \leq T$ for $i < j$.

Let $\widehat{\ }$ and $\underbrace{\ }$ be the shorthands for sup and inf operator respectively. Since the mass function of NPI for precise Bernoulli data assign only assign non zero value to consecutive integer set in $\mathscr{C}(\mathbb{N}_0^T)$, if one rewrite $S_T \in \{z_i\}_{i=1}^{i=\alpha}$ as $S_T \in \bigcup\limits_{1\leq h\leq l} R_h$ where each R_h, $h \in \mathbb{N}_1^l$ is a consecutive integer set and $\widehat{R_h}+1 < \underbrace{R_{h+1}}$. For example, $\{z_i\}_{i=1}^{i=\alpha} = \{4,7,9,10\} = \bigcup\limits_{1\leq h\leq 3} R_h = \{4\} \cup \{7\} \cup \{9,10\}$, then:

$$
\begin{aligned}
\underline{p}_{(n,j)}(S_T \in \{z_i\}_{i=1}^{i=\alpha}) &= \underline{p}_{(n,j)}(S_T \in \underset{1 \le h \le l}{\cup} R_h) \\
&= \sum_{h=1}^{l} \underline{p}_{(n,j)}(S_T \in R_h) \\
&= \sum_{h=1}^{l} \sum_{a \in \mathscr{C}(R_h)} m_{(n,j)}(S_T \in a) \\
&= \binom{n+T}{n}^{-1} \times \sum_{h=1}^{l} \sum_{\substack{r_1, r_2 \in R_h \\ r_1 \le r_2}} \left[\binom{j-1+r_1}{r_1} \binom{n-j-1+T-r_2}{T-r_2} \right]
\end{aligned}
$$

3.2.3 Consistence of Mass function in NPI for Bernoulli data

One can show that the mass function in the NPI for Bernoulli data is consistent as defined in Definition 2.1.5.

Given data (n, j), NPI for Bernoulli data induce a mass function on $[\Omega = \{B, B^c\}^T, \mathscr{A}_T = \sigma(S_T^{-1})]$ $\forall T \in \mathbb{N}$. Let $\mathcal{P}$ denote the power set operator then:

$$\mathscr{A}_T = \{S_T^{-1}(\epsilon) | \epsilon \in \mathcal{P}(\mathbb{N}_0^T)\}$$

Define a binary operator $\ominus : \mathcal{P}(\mathbb{N}_0^T) \times \mathbb{N}_0^T \longrightarrow \mathcal{P}(\mathbb{N}_0^T)$ as following:

For $A \in \mathcal{P}(\mathbb{N}_0^T)$ and $b \in \mathbb{N}_0^T$

$$A \ominus b = \underset{a \in A}{\cup} \underset{j \in \mathbb{N}_0^b}{\cup} \{max(0, a - j)\}$$

For example, $\{1, 5, 9\} \ominus 2 = \{0, 1, 3, 4, 5, 7, 8, 9\}$

For $[\Omega, \mathscr{A}_i, m^i_{(n,j)}]$, $i \in \mathbb{Z}_1^T$. The case that "$\epsilon_k \in \mathscr{A}_k$, $\epsilon_l \in \mathscr{A}_l$, $k \neq l$, $\epsilon_k \subset \epsilon_l$" could happen in NPI for Bernoulli data is when

$$
\begin{aligned}
\mathscr{A}_k &= \{S_k^{-1}(\epsilon) | \epsilon \in \mathcal{P}(\mathbb{Z}_0^k)\} \quad , \quad S_k = \sum_{i_1=1}^{k} X_{r_{i_1}} \\
\mathscr{A}_l &= \{S_l^{-1}(\epsilon) | \epsilon \in \mathcal{P}(\mathbb{Z}_0^l)\} \quad , \quad S_l = \sum_{i_2=1}^{l} X_{r_{i_2}}
\end{aligned}
$$

$l < k$ and $S_k \in \epsilon_1$ for some $\epsilon_1 \in \mathcal{P}(\mathbb{Z}_0^k)$ and $\exists \epsilon_2 \in \mathcal{P}(\mathbb{N}_0^l)$ such that $\epsilon_1 \ominus (k - l) \subset \epsilon_2$

then $\epsilon_k \subset \epsilon_l$ where $\epsilon_k = \{\omega | S_k(\omega) \in \epsilon_1\}$ and $\epsilon_l = \{\omega | S_l(\omega) \in \epsilon_2\}$. Since NPI Bernoulli lower probability has the property:

$$\underline{p}_{(n,j)}(S_T \in \bigcup_{1 \leq h \leq l} R_h) = \sum_{h=1}^{l} \underline{p}_{(n,j)}(S_T \in R_h)$$

where R_h are different consecutive integer blocks, one only need show the case $\epsilon_1 = \mathbb{N}_{m_1}^{m_2}$ and $\epsilon_2 = \mathbb{N}_{m_1-(k-l)}^{m_2}$ has the following:

$$\underline{p}_{(n,j)}(S_l \in \mathbb{N}_{m_1-(k-l)}^{m_2}) \geq \underline{p}_{(n,j)}(S_k \in \mathbb{N}_{m_1}^{m_2}) \quad \forall (l-k) \in \mathbb{Z}$$

And this is true if:

$$\underline{p}_{(n,j)}(S_{k-1} \in \mathbb{N}_{m_1-1}^{m_2}) \geq \underline{p}_{(n,j)}(S_k \in \mathbb{N}_{m_1}^{m_2})$$

which is followed by Formulae 2.2.28

3.3 Construction of imprecise expectation measures for a general function of S_T

With the mass function in NPI for Bernoulli data, one now can use the proposed general algorithm in Section 3.1 to construct imprecise expectation measures for a general function of S_T and hence enable the computation of imprecise expectations of a general function of S_T. An example is provided below.

Example: Consider $[\Omega = \{B, B^c\}^5, \mathscr{A}_5 = \sigma(S_5^{-1}), m_{(n,j)}(S_5)]$, given Bernoulli data (n, j) and a function f of S_5 induce a order $\succ$ on $\mathbb{N}_0^5$ which form a partition $I = \{\{2,4\}, \{1,3\}, \{0,5\}\}$ of $\mathbb{N}_0^5$ with order as they appear. So $\{2,4\} \succ \{1,3\} \succ \{0,5\}$ and there are ties between 2 and 4, 1 and 3, also 0 and 5.

Since the NPI induced Bernoulli imprecise space only have non zero mass on the set $\mathscr{C}(\mathbb{N}_0^5)$, one now can initiate the algorithm with $Q_y^3 = \mathscr{C}(\mathbb{N}_0^5)$ instead of $\mathcal{P}(\mathbb{N}_0^5)$, which could reduce some steps in the mass assignment.

Apply GMA algorithm, start with $Q_0 = (Q_0^1, Q_0^2, Q_0^3, Q_0^4) = (\mathbb{N}_0^5, \{2,4\}, \mathscr{C}(\mathbb{N}_0^5), \emptyset)$

One assigns all probability mass of the elements in $Q_0^3 = \mathscr{C}(\mathbb{N}_0^5)$ which have non empty intersection with $\{2,4\}$ to $\{2,4\}$. Denote the set of those elements as G_1, then $G_1 = \{\{0,1,2,3,4,5\},\{0,1,2,3,4\},\{1,2,3,4,5\},\{0,1,2,3\},\{1,2,3,4\},\{2,3,4,5\},$ $\{0,1,2\},\{1,2,3\},\{2,3,4\},\{3,4,5\},\{1,2\},\{2,3\},\{3,4\},\{4,5\},\{2\},\{4\}\}$, And denote the probability mass value assign to $\{2,4\}$ as p_1. Then $p_1 = \sum_{\epsilon \in G_1} m_{(n,j)}(S_5 \in \epsilon)$.

Now one moves to the next stage $Q_1 = (Q_1^1, Q_1^2, Q_1^3, Q_1^4)$, where $Q_1^1 = \{0,1,3,5\}$, $Q_1^2 = \{1,3\}$, $Q_1^3 = \mathscr{C}(\mathbb{Z}_0^5) \setminus G_1$, $Q_1^4 = \{(\{2,4\},p_1)\}$.

One now starts to assign probability mass to second order element, in this case is $Q_1^2 = \{1,3\}$. Denote the set of elements in Q_1^3 which have non empty intersection with {1,3} as G_2, then $G_2 = \{\{0,1\},\{1\},\{3\}\}$. Also denote the probability mass value assigned to $\{1,3\}$ as p_2, then $p_2 = \sum_{\epsilon \in G_2} m_{(n,j)}(S_5 \in \epsilon)$ to $\{1,3\}$.

Now one can move to stage $Q_2 = (Q_2^1, Q_2^2, Q_2^3, Q_2^4)$, where $Q_2^1 = \{0,5\}$, $Q_2^2 = \{0,5\}$, $Q_2^3 = Q_1^3 \setminus G_2$, $Q_2^4 = \{(\{2,4\},p_1),(\{1,3\},p_2)\}$.

Since $Q_2^1 \neq \emptyset$, one now still needs to move to the next stage, one assigns probability mass to the last order element, which is $Q_2^2 = \{0,5\}$. The residual elements in Q_2^3 are $\{\{0\},\{5\}\}$ and both of them have non empty intersection with {0,5}. Denote the set of the residual elements in Q_2^3 as G_3, and probability mass value assigned to $\{0,5\}$ as p_3, then $p_3 = \sum_{\epsilon \in G_3} m_{(n,j)}(S_5 \in \epsilon)$.

Now move to stage $Q_3 = (\emptyset, \emptyset, \emptyset, \{(\{2,4\},p_1),(\{1,3\},p_2),(\{0,5\},p_3)\})$ and stop.

In total, $\frac{(1+5)(2+5)}{2} = 21$ mass atoms are assigned and distributed into $\{p_1, p_2, p_3\}$ and Q_3^4 is the upper expectation measure for $f(S_5)$. To find the lower expectation measure of $f(S_5)$, one simply needs to use the reverse order of $\succ$.

3.4 NPI for imprecise Bernoulli data

NPI for imprecise Bernoulli data was firstly considered by Coolen in 2008 [16]. Coolen presented the condition reasoning in detail. However, the attention was limited to the set-value data and the computation of imprecise probability remains unsolved. By extending the path counting concept in the lattice representation of NPI for Bernoulli data and considering the most general form of imprecise Bernoulli data, we further develop NPI for imprecise Bernoulli data in this section.

Let us firstly recall the precise data mass function Formula 3.2.1 below:

$$m_{(n,j)}(S_T) = \begin{cases} \binom{j-1+r_1}{r_1}\binom{n-j-1+T-r_2}{T-r_2} \times \binom{n+T}{n}^{-1} & \exists \epsilon \in \mathscr{C}(\mathbb{N}_0^T) \text{ s.t. } S_T \in \epsilon = \mathbb{N}_{r_1}^{r_2} \\ 0 & \nexists \epsilon \in \mathscr{C}(\mathbb{N}_0^T) \text{ s.t. } S_T \in \epsilon = \mathbb{N}_{r_1}^{r_2} \end{cases}$$

and also its corresponding path counting Figure 3.3

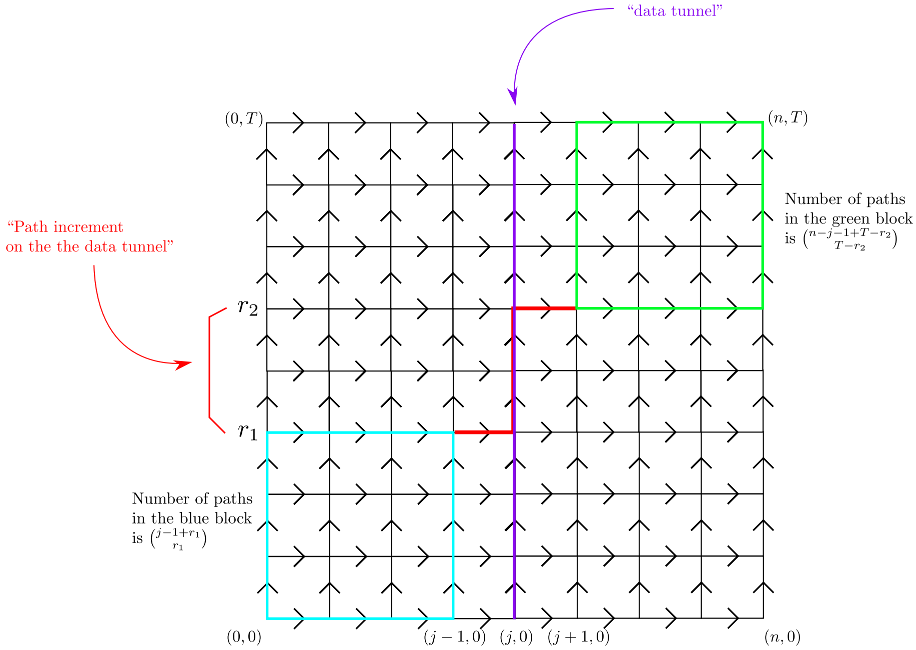

Figure 3.6: Figure 3.3 with new description

Given data (n, j), if we defined $x = j$ as a data tunnel, indicated by the purple line in Figure 3.6 as "data tunnel", $S_T \in \mathbb{N}_{r_1}^{r_2}$ as inference target, path upward movement in the y direction as increment, then legitimate paths in the precise data mass function Formula 3.2.1 are the paths of which the increment on the data tunnel covers exactly the inference target. If the inference target S_T belongs to some non consecutive number sets, then the data tunnel will not be able to cover the inference target, therefore no legitimate path exists. This explains why when $\nexists \epsilon \in \mathscr{B}_T$ s.t $S_T \in \epsilon$, the precise data mass function Formula 3.2.1 yields zero value.

Analogously, given imprecise Bernoulli data which is the form of several intervals, one can extend the legitimate paths concept and constructs the mass function for imprecise Bernoulli data.

3.4.1 The mass function and its corresponding lattice path counting

Given a set $D_{(n)} = \{x_i\}_{i=-n+1}^{0}$ of imprecise Bernoulli data with $x_i \in \{0,1\}$ and $J = |\{i : x_i = 1 \wedge x_i \in D_{(n)}\}| = \{j_1, j_2, j_3, ..., j_e\}$, $e \leq n$, denoted the imprecise data as (n, J), one could rewrite the set J as $J = \underset{1\leq i\leq k}{\cup} J_i$, $k \leq n$ where each $J_i = \mathbb{N}_a^b$ for some a,b, and $\widehat{J_i} + 1 < \underset{\smile}{J_{i+1}}$, $\forall i$
Still consider the random variable S_T defined before and rewrite each event as $S_T \in \underset{1\leq h\leq l}{\cup} R_h$, $l < T$ where each R_i is set with same form of J_i.

Define $J_0 = \{-1\}$, $J_{k+1} = \{n+1\}$, $R_0 = \{0\}$, $R_{l+1} = \{T\}$

Then NPI mass function for $S_T \in \underset{1\leq h\leq l}{\cup} R_h$ is:

$$
\begin{aligned}
& m_{(n,J=\underset{1\leq i\leq k}{\cup} J_i)}(S_T \in \underset{1\leq h\leq l}{\cup} R_h) \qquad (3.4.2)\\
= & \binom{n+T}{n}^{-1} \times \sum_{C} \prod_{i=1}^{k+1} \left[\binom{\underset{\smile}{J_i} - \widehat{J_{i-1}} - 2 + r_{2i-1} - r_{2i-2}}{\underset{\smile}{J_i} - \widehat{J_{i-1}} - 2} \binom{\widehat{J_{i-1}} - \underset{\smile}{J_{i-1}} + r_{2i-2} - r_{2i-3}}{\widehat{J_{i-1}} - \underset{\smile}{J_{i-1}}} \right]
\end{aligned}
$$

where $C = \{\{r_i\}_{-1}^{2k+1} | r_0 \in R_0$ and $r_{-1} \in R_0$; $r_1 = \underset{\smile}{R_1}$; $r_{2k} = \widehat{R_l}$; $r_{2k+1} \in R_{l+1}$; $\forall n \in \mathbb{N}_0^{k-1}, r_{1+2n} \in R_h$ and $r_{2+2n} \in R_h$ for some $h \in \mathbb{N}_1^l$; $r_x \leq r_y, \forall x \leq y$; let $R'_i = \mathbb{N}_{r_{2i-1}}^{r_{2i}}$ for $i \in \mathbb{N}_1^k$, then $\underset{i\in\mathbb{N}_1^k}{\cup} R'_i = \underset{i\in\mathbb{N}_1^l}{\cup} R_i\}$.

In the case $l > k$, $C = \emptyset$ and the mass function yields zero value.

When $J = \mathbb{N}_a^b$, the above formula could be simplified to:

$$
\begin{aligned}
& m_{(n,J=\mathbb{N}_a^b)}(S_T \in \{R_1\})\\
= & \binom{n+T}{n}^{-1} \times \left[\binom{a-1+\underset{\smile}{R_1}}{a-1} \binom{b-a+\widehat{R_1}-\underset{\smile}{R_1}}{b-a} \binom{n-b-1+T-\widehat{R_1}}{n-b-1} \right]
\end{aligned}
$$

The summing logic C of the NPI imprecise data mass function Formula 3.4.2, in essence, has the same notion of legitimate path counting in the NPI lattice representation graph. The only modification is that, given data (n, J) with k consecutive number blocks $J = \underset{1 \le i \le k}{\cup} J_i$, we now have k units of data tunnels and they are thicker than the precise case. This results in non zero mass value for inference target $S_T \in \underset{1 \le h \le l}{\cup} R_h$ with $l \le k$ consecutive number blocks.

We illustrate imprecise data legitimate paths counting with a small example when $k = 3$ in the data, in other words, $J \in \underset{1 \le i \le 3}{\cup} J_i$ in below.

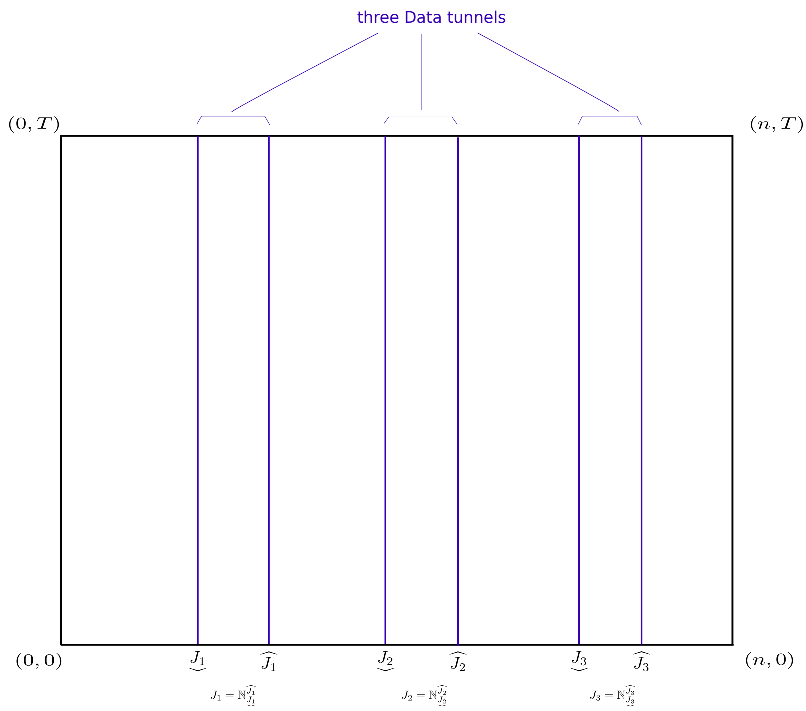

Figure 3.7: Data tunnels in imprecise data path counting

Figure 3.7 shows the data tunnels in the imprecise data case. To be a legitimate path fitting the criteria that the increment on the data tunnels covers exactly the inference target, depending on the inference target, there are only certain numbers of entrance and exit points for the path to access the data tunnels.

Let us firstly consider reference target $l = 2 < k$, $S_T \in \underset{1 \le h \le 2}{\cup} R_h$

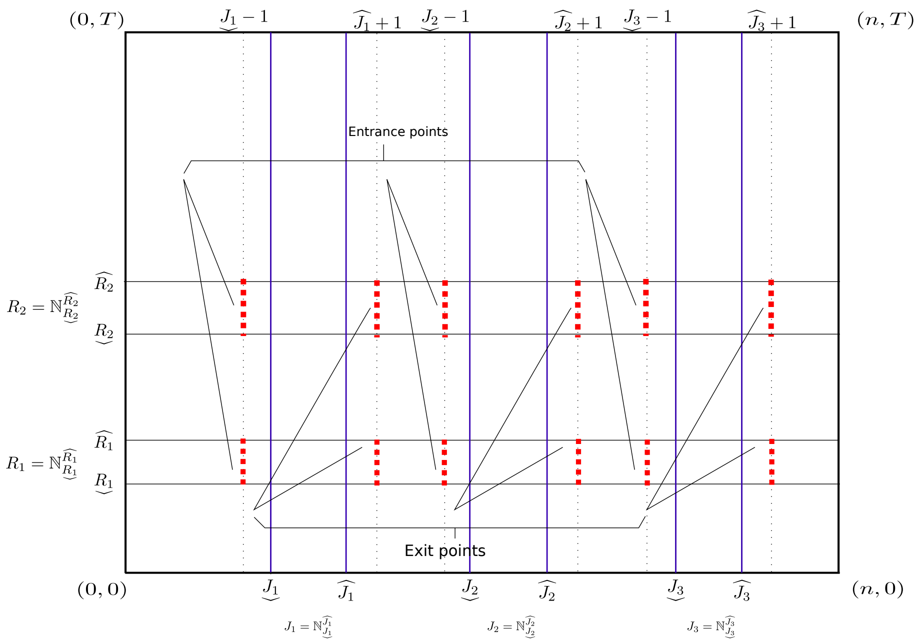

Figure 3.8: Legitimate entrance and exit points for each data tunnels when $J \in \bigcup_{1\leq i\leq 3} J_i$ and $S_T \in \bigcup_{1\leq i\leq 2} R_h$

In this case, the legitimate entrance and exit points for each data tunnels are illustrated in Figure 3.8. Due to the existence of legitimate entrance and exit points and the paths are only allowed to going upward or rightward, one knows that the maximum cover range for one data tunnels is one block R_i. So when $l < k$, namely the number of blocks in the inference target S_T is less than the number of blocks in the data J, there could be some slackness in the usage of the data tunnels. Therefore the legitimate paths are allowed to partially cover the some of inference blocks in some data tunnels or skip using some of the data tunnels. Using the example above, we illustrate this in Figure 3.9 and Figure 3.10.

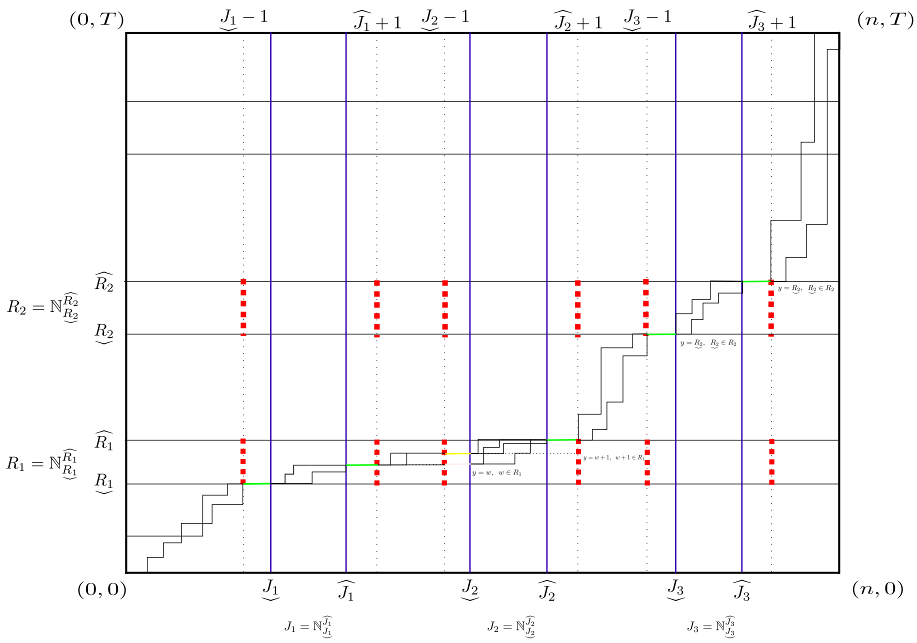

Figure 3.9: Example of paths that partially cover one of inference block in one data tunnel

In Figure 3.9, the paths only partially cover first the inference target of S_T using the first data tunnel J_1, in order to be legitimate, these paths must end at point (n, T) passing through all the green line and one of yellow or pink line.

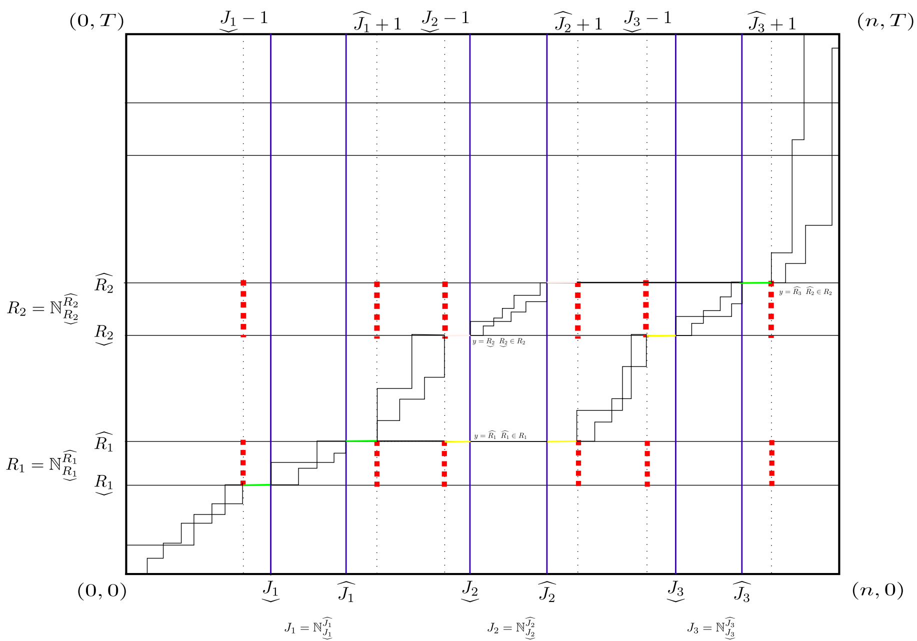

Figure 3.10: Example of paths that skipping using one of data tunnels

In Figure 3.10, the paths could skip using the second data tunnel J_2 by passing through all the green line and yellow line or skipping using the third data tunnel J_3 by passing through all the green line and pink line.

When $l = k$, namely the number of block in the inference target S_T is equal the number of block in the data J, there is no slackness in the usage of the data tunnels. Thus, all the legitimate paths have to fully use each data tunnel to cover one inference target block. Using the same example, we demonstrate this in Figure 3.11.

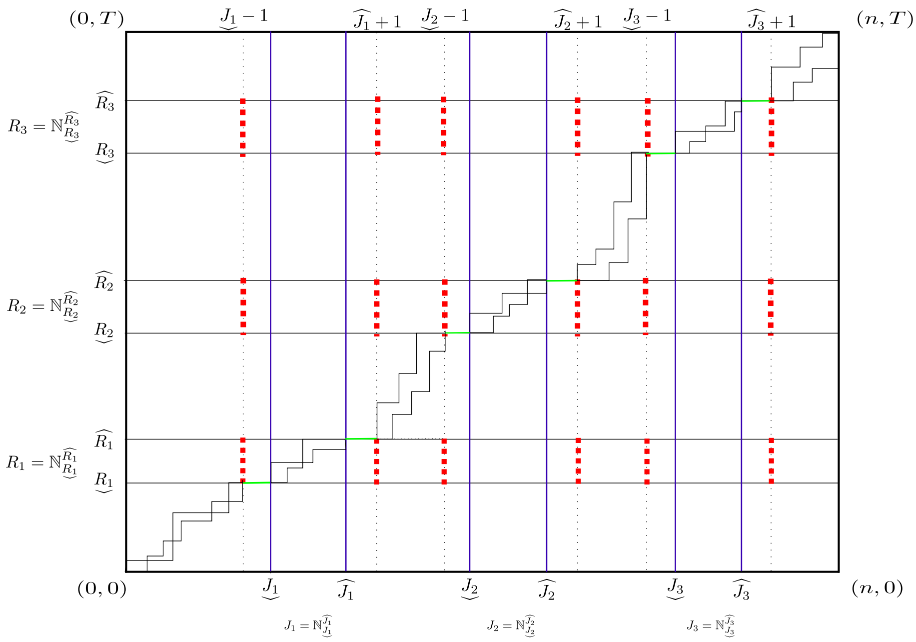

Figure 3.11: Example of paths which fully use each data tunnels when the number of blocks in the data and the inference target are equal.

In Figure 3.11, since $k = l = 3$, the paths have to fully use each data tunnel J_i to cover each inference target R_i for all $i \in \{1, 2, 3\}$ by passing through all the green line.

Since the maximum inference target cover range for one data tunnels is one block R_i, when $l > k$, the number of blocks in the inference target S_T is greater the number of blocks in the data J, there is not enough number of data tunnels to cover all the reference target blocks. Therefore, no legitimate path exists in this case.

3.4.2 The imprecise probability of NPI for imprecise Bernoulli data

Before calculating the lower and upper probability for NPI imprecise Bernoulli data, it is useful to see what subsets in the sample space Ω have non zero mass value in the mass function.

Denote $\lfloor\cdot\rfloor$ as the floor function. $\lfloor x\rfloor = \max\{z|\ z \in \mathbb{Z} \text{ and } z \leq x\}$

Let us firstly recall the operator $\mathscr{C}$ defined previously and extend it to $\mathscr{C}_e$, $e \in \mathbb{N}^+$

$$\mathscr{C}(\mathbb{N}_{i_1}^{i_2}) = \{\mathbb{N}_{j_1}^{j_2}|j_1, j_2 \in \mathbb{N}_{i_1}^{i_2}, i_1 \leq j_1 \leq j_2 \leq i_2\}$$

$$\mathscr{C}_e(\mathbb{N}_{i_1}^{i_2}) = \{\bigcup_y \mathbb{N}_{j_{(y,1)}}^{j_{(y,2)}}|j_{(y,1)}, j_{(y,2)} \in \mathbb{N}_{i_1}^{i_2},\ j_{(y,1)} \leq j_{(y,2)},\ j_{(y,2)} < j_{(y+1,1)} + 1\ \forall y \in \mathbb{N}_1^e\}$$

$\mathscr{C}_e$ will generates a collection of subsets which contain e disjoint blocks consecutive positve values of the argument $\mathbb{N}_{i_1}^{i_2}$. One could know that the previous defined $\mathscr{C} = \mathscr{C}_1$.

Let k be the number of consecutive integer blocks in the data, l be number of consecutive integer blocks in the reference target. One could know for S_T, the maximum number for l is $\lfloor \frac{T}{2} \rfloor + 1$. Let $k^* = \min(k, \lfloor \frac{T}{2} \rfloor + 1)$, then the set $\mathscr{B}_T^k = \{S_T^{-1}(a) \mid a \in \bigcup_{e \in \mathbb{N}_1^{k^*}} \mathscr{C}_e(\mathbb{N}_0^T)\}$ contains all the subsets of Ω which has non zero mass value. This comes from the fact that when the number of blocks l in the reference target exceeds the number of blocks k in the data, the mass function yields zero value.

The NPI lower probability for $S_T \in \bigcup_{1 \leq h \leq l} R_h$ is then:

$$
\begin{aligned}
& \underline{p}_{(n,J=\bigcup\limits_{1\le i\le k} J_i)}(S_T \in \bigcup_{1\le h\le l} R_h) \\
= & \sum_{\substack{\epsilon\in\mathcal{P}(\mathbb{N}_0^T) \\ \epsilon\subset \bigcup\limits_{1\le h\le l} R_h}} m_{(n,J)}(S_T\in\epsilon) && (3.4.3) \\
= & \sum_{\substack{\epsilon\in \bigcup\limits_{e\in\mathbb{N}_1^{\lfloor k^*\rfloor}} \mathscr{C}_e(\mathbb{N}_0^T) \\ \epsilon\subset \bigcup\limits_{1\le h\le l} R_h}} m_{(n,J)}(S_T\in\epsilon) && (3.4.4) \\
= & \binom{n+T}{n}^{-1} \times \sum_{\underline{C}} \prod_{i=1}^{k+1} \left[\binom{\underset{\smile}{J_i} - \widehat{J_{i-1}} - 2 + r_{2i-1} - r_{2i-2}}{\underset{\smile}{J_i} - \widehat{J_{i-1}} - 2} \binom{\widehat{J_{i-1}} - \underset{\smile}{J_{i-1}} + r_{2i-2} - r_{2i-3}}{\widehat{J_{i-1}} - \underset{\smile}{J_{i-1}}} \right] && (3.4.5)
\end{aligned}
$$

where $\underline{C} = \{\{r_i\}_{-1}^{2k+1} | r_0 \in R_0 \text{ and } r_{-1} \in R_0\ ;\ r_{2k+1} \in R_{l+1}\ ;\ \forall n \in \mathbb{N}_0^{k-1}, r_{1+2n} \in R_h \text{ and } r_{2+2n} \in R_h \text{ for some } h \in \mathbb{Z}_1^l\ ;\ r_x \le r_y, \forall x \le y\}$.

Equality 3.4.5 comes from legitimate paths for the NPI lower probability counting in the lattice representation. From Equation 3.4.3, one could know the legitimate paths in the lower probability are the paths of which the increment on the data tunnels only covers any subset of the inference target. In other words, using example $J \in \bigcup\limits_{1\le i\le 3} J_i$ and $S_T \in \bigcup\limits_{1\le h\le 3} R_h$, the lower probability is total number of paths that within each data tunnels only use any subset of the entrance and exit points to reach (n,T) from $(0,0)$ in Figure 3.11.

When $J = \mathbb{N}_a^b$, the lower probability formula for imprecise data could be simplified to:

$$
\begin{aligned}
& \underline{p}_{(n,J=\mathbb{N}_a^b)}(S_T \in \bigcup_{1\le h\le l} R_h) \\
= & \sum_{1\le i<\le l} \sum_{\substack{r_{(i,1)}, r_{(i,2)} \in R_i \\ r_{(i,2)} > r_{(i,1)}}} \left[\binom{a-1+r_{(i,1)}}{a-1} \binom{b-a+r_{(i,2)}-r_{(i,1)}}{b-a} \binom{n-b-1+T-r_{(i,2)}}{n-b-1} \right] \\
& \times \binom{n+T}{n}^{-1} && (3.4.6)
\end{aligned}
$$

Moreover, consider $S_T \in \mathbb{N}_a^b$, let c,d be the integer numbers and $0 \le c \le d \le n$, the lower probability formula for imprecise data has following equality:

$$\underline{p}_{(n,J_1=\epsilon_1\cup\{c,d\})}(S_T \in \mathbb{N}_a^b) = \underline{p}_{(n,J_2=\epsilon_2\cup\{c,d\})}(S_T \in \mathbb{N}_a^b) \quad \forall\epsilon_1,\epsilon_2 \subset \mathbb{N}_c^d \tag{3.4.7}$$

This equality comes from the fact that when the reference $S_T \in \mathbb{N}_a^b$ is one block of consecutive integer number, the lower probability path counting for $J = \{c,d\}$ is exactly the same lower probability path counting for $J' = \{c,d\} \cup \epsilon'$ with $\epsilon' \subset \mathbb{N}_c^d$. Figure 3.12 gives one graphical example.

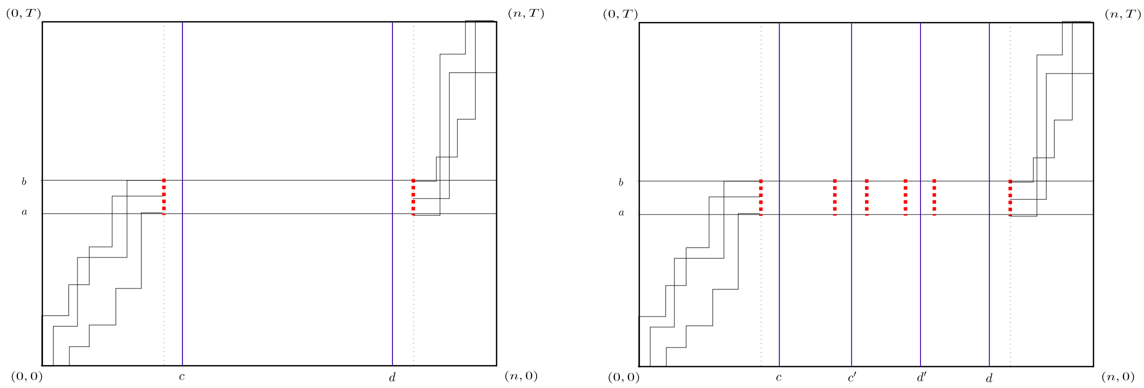

Figure 3.12: $J_1 = \emptyset \cup \{c,d\}$ and $J_2 = \{c',d'\} \cup \{c,d\}$ have the same lower probability counting for $S_T \in \mathbb{N}_a^b$

In Figure 3.12, given reference target $S_T \in \mathbb{N}_a^b$, $c < c' < d' < d$, the data $J_1 = \{c,d\}$ and data $J_2 = \{c,c',d',d\}$ have the same legitimate paths counting in the lower probability formula, the legitimate paths of the lower probability in both case are the paths enter any left-most entrance points and exit from any right-most exit point. With the imprecise probability conjugacy property $\overline{p}(A^c) = 1 - \underline{p}(A)$, one also has:

$$\overline{p}_{(n,J_1=\epsilon_1\cup[0,c)\cup(d,n])}(S_T \in \mathbb{N}_a^b) = \overline{p}_{(n,J_2=\epsilon_2\cup[0,c)\cup(d,n])}(S_T \in \mathbb{N}_a^b) \quad \forall\epsilon_1,\epsilon_2 \subset \mathbb{N}_{c-1}^{d-1} \tag{3.4.8}$$

Using the path counting argument, one also has the NPI upper probability for imprecise Bernoulli data. The NPI upper probability for imprecise Bernoulli data of $S_T \in \underset{1\leq h\leq l}{\cup} R_h$ is:

$$\begin{aligned}
&\overline{p}_{(n,J=\underset{1\leq i\leq k}{\cup} J_i)}(S_T \in \underset{1\leq h\leq l}{\cup} R_h) \\
&= \sum_{\substack{\epsilon\in\mathcal{P}(\mathbb{N}_0^T) \\ \epsilon\cap \underset{1\leq h\leq l}{\cup} R_h\neq\emptyset}} m_{(n,J)}(S_T\in\epsilon) && (3.4.9) \\
&= \sum_{\substack{\epsilon\in \underset{e\in\mathbb{N}_1^{\lfloor k^*\rfloor}}{\cup} \mathscr{C}_e(\mathbb{N}_0^T) \\ \epsilon\cap \underset{1\leq h\leq l}{\cup} R_h\neq\emptyset}} m_{(n,J)}(S_T\in\epsilon) && (3.4.10) \\
&= \binom{n+T}{n}^{-1} \times \sum_{\overline{C}} \prod_{i=1}^{k+1} \left[\binom{\underbrace{J_i} - \widehat{J_{i-1}} - 2 + r_{2i-1} - r_{2i-2}}{\underbrace{J_i} - \widehat{J_{i-1}} - 2} \binom{\widehat{J_{i-1}} - \underbrace{J_{i-1}} + r_{2i-2} - r_{2i-3}}{\widehat{J_{i-1}} - \underbrace{J_{i-1}}} \right] && (3.4.11)
\end{aligned}$$

where $\overline{C} = \{\{r_i\}_{-1}^{2k+1} | r_0 \in R_0 \text{ and } r_{-1} \in R_0 \ ; \ r_{2k+1} \in R_{l+1} \ ; \ \forall i \ , \ r_i \in \mathbb{N}_0^T \ ; \ \forall n \in \mathbb{N}_1^k, \ \forall h \in \mathbb{Z}_1^l, \ \mathbb{N}_{r_{2n-1}}^{r_{2n}} \cap R_h \neq \emptyset \ ; \ r_x \leq r_y, \forall x \leq y \ \}$.

The legitimate paths in the NPI upper probability for imprecise data are the paths of which the increment on any data tunnels contains any subset of the inference target.

To find the imprecise expectation measure of $f(S_T)$ in NPI for imprecise Bernoulli data, one simply needs to use algorithm in Section 3.1 and initiate it with $Q_0^3 = \underset{e\in\mathbb{N}_1^{\lfloor k^*\rfloor}}{\cup} \mathscr{C}_e(\mathbb{N}_0^T)$.

3.4.3 Property of NPI imprecise probability for Bernoulli imprecise data

In this section, we present the property of NPI for imprecise Bernoulli data, followed by a numerical example. This property is then explained by the lattice representation path counting argument.

Property Given data (n, J) and reference target $S_T \in \epsilon$. The imprecision in the imprecise probability increase as the imprecision in the data J increases, more precisely one has:

$$\underline{p}_{(n,J_1)}(S_T \in \epsilon) \geq \underline{p}_{(n,J_2)}(S_T \in \epsilon) \quad \text{if} \quad J_1 \subset J_2 \tag{3.4.12}$$

$$\overline{p}_{(n,J_1)}(S_T \in \epsilon) \leq \overline{p}_{(n,J_2)}(S_T \in \epsilon) \quad \text{if} \quad J_1 \subset J_2 \tag{3.4.13}$$

Table 3.1: Numerical examples for the property

Value of $[\underline{p}_{(n,J)}, \overline{p}_{(n,J)}]$		
Reference target S_T / Data (n, J)	$S_{10} \in \{4, 5, 6\}$	$S_{13} \in \{6, 7, 9\}$
(10, (5,6))	[0.2398 , 0.7256]	[0.0990 , 0.7499]
(10, (5,6,7))	[0.1306 , 0.7802]	[0.0395 , 0.8481]
(10, (1,5,6,7))	[0.0040 , 0.8517]	[0.0005 , 0.8755]
(10, (1,3,5,6,7))	[0.0040 , 0.9387]	[0.0003 , 0.9259]
(10, $\mathbb{N}_0^{10}$)	[0 , 1]	[0 , 1]

From Table 3.1, notice that under Data (10,(1,5,6,7)) and (10,(1,3,5,6,7)), the reference target $S_{10} \in \{4, 5, 6\}$ have the same lower probability value 0.0040 which give a numerical example of the Equality 3.4.7. One could also observe that given 10 data point, with increasing imprecise in the data, the gap between the lower and upper probability value induced by NPI for imprecise Bernoulli data become wider as it is stated in Inequality 3.4.12 and 3.4.13.

To understand why does this property exist, one should know how the increase of imprecision in data affects the lower probability legitimate path counting.

Consider example $J \in \underset{1 \leq i \leq 2}{\cup} J_i$ and $S_T \in \underset{1 \leq h \leq 2}{\cup} R_h$. Recall that legitimate paths in the lower probability are the paths of which the increment on the data tunnels covers any subset of the inference target, legitimate paths in the example are then paths that have to be confined in green area within each data tunnel. (See Figure 3.13)

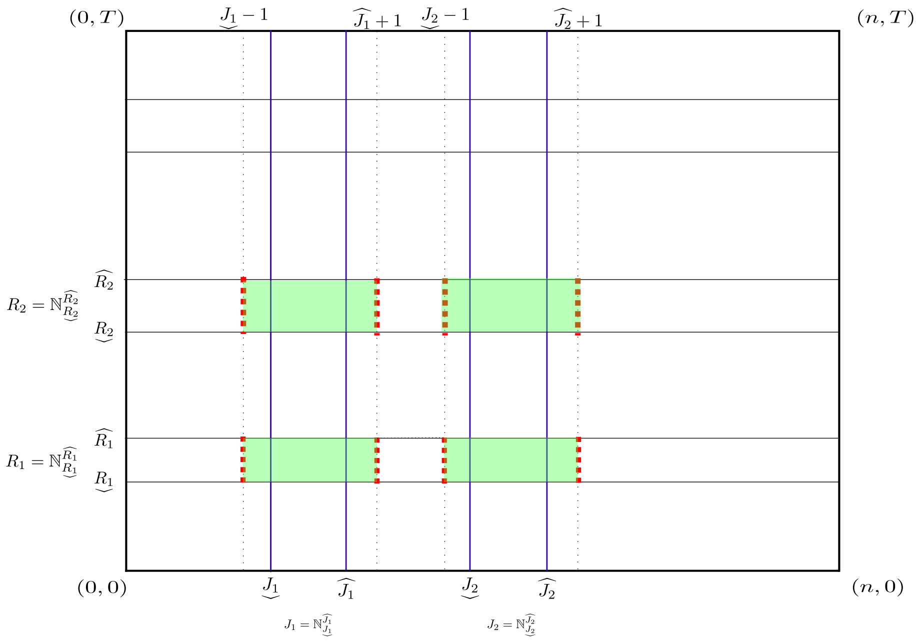

Figure 3.13: Graphical illustration of property in NPI for imprecise Bernoulli data

These green areas could be thought of the restrictions for the paths that legitimately go from $(0,0)$ to (n,T). The increment of imprecision in the data will either expand the width of existing data tunnel or create a new data tunnel, these two results will both lead to an increment of total size of green areas within the lattice and thus reduce the number of legitimate paths in the lower probability mass counting.

The above argument gives the reason for Inequality 3.4.12. Since increasing imprecision in data leads to decrement in the lower probabilities of all reference target and from the imprecise probability conjugacy property 2.1.2, one knows $\overline{p}(A) = 1 - \underline{p}(A^c)$, the upper probability value will increase as imprecision in data increase. This gives the reason for Inequality 3.4.13.

Chapter 4

Application of NPI method in asset trading

In this chapter, under binomial tree model, considering the financial object asset A_T defined in Chapter 2, we apply NPI to learn the information from historical data, induce an imprecise probability space on the asset price $A_T(S_T)$ and study the performance of two different NPI asset trading routes in a simple scenario setting.

4.1 Asset Scenario setting

Consider the scenario: one is allowed to long or short the one unit of asset at price a_0 at time 0. Additionally one is allowed to invest or borrow a_0 with risk free interest rate r. Whatever position one enters, one has to keep the position for time length T and is obligated to close all his risk position at time T, how should one, who is a NPI imprecise probability believer, without using any of his or her capital, make one's decision in trading to maximize one's capital gain in present value probabilistically or expectationally at time T? (Assume one's capital is able to cover any potential loss)

In the scenario, the key points to emphasize are: fixed entering position time point, fixed closing position time point, one single asset is available to long or short.

One is interested in the asset price $A_T = a_0 u^{S_T} d^{T-S_T}$ at time T. Since $A_T(S_T)$ is a monotonically increasing and non-negative function of S_T, one is able construct

$p_{\overline{E}(A_T)}$ and $p_{\underline{E}(A_T)}$ by Formulas 2.2.30 and 2.2.29. Using $p_{\overline{E}(A_T)}$ and $p_{\underline{E}(A_T)}$, one then is able to find $\underline{E}(A_T)$ and $\overline{E}(A_T)$

A NPI believer, who prefer to use imprecise probability operator $\underline{p}$ and $\overline{p}$ could rationally make following decision on asset trading in this scenario at time 0:

Set threshold value $0.5 < w < 1$. From NPI setting, one could know $0 < \underline{p}(A_T(S_T) > a_0B(T)) < \overline{p}(A_T(S_T) > a_0B(T)) < 1$ if $S_T \subsetneq \mathbb{N}_0^T$ and $S_T \neq \emptyset$.

Imprecise probability asset trading route 1.1:

$$
\begin{cases}
\text{Borrow cash } a_0 \text{ and buy the asset } a_0 \text{ at time 0. And sell the asset at time } T \\
\text{return } a_0B(T) \text{ to the lender. if } \underline{p}(A_T(S_T) > a_0B(T)) > w \\
\\
\text{Short the asset at for } a_0 \text{ current time, invest the cash } a_0 \text{ at risk free rate and} \\
\text{close the both positions at time } T \text{ if } \underline{p}(A_T(S_T) < a_0B(T)) > w \\
\\
\text{No action if none of above satisfied}
\end{cases}
$$

Motivation behind NPI imprecise probability asset trading route 1.1:

When $A_T(S_T) > a_0B(T)$, the increment $A_T - a_0$ of asset price from time 0 to time T is greater than the interest $a_0 \times (e^r - 1)$ generated at time T from borrowing a_0 at time 0. So when the lower probability of this event $A_T(S_T) > a_0B(T)$ is greater than the threshold value w, a prudent individual using imprecise probability will invest in the asset a_0 and borrow a_0, anticipating the profit $A_T(S_T) - a_0B(T)$ at time T by closing the position.

On the contrary, the event $A_T(S_T) < a_0B(T)$ indicates the increment $A_T - a_0$ of the asset price from time 0 to time T is less than the interest $a_0 \times (e^r - 1)$ generated at time T by investing a_0 in risk free rate r at time 0. If the lower probability of event $A_T(S_T) < a_0B(T)$ is greater than the threshold value w, then one would better off short sell the asset for a_0 and invest a_0 in risk free rate r, anticipating the profit $a_0B(T) - A_T(S_T)$ at time T by closing the position.

If none of the above conditions satisfied, one would better do nothing as the information learned from historical data is not "certain" enough for one to make a confident trade.

One can show that only one of the actions could be taken in route 1.1: Using inequality and conjugacy property of imprecise probability, one could know:

$$
\begin{aligned}
&\text{if } \underline{p}(A_T(S_T) > a_0B(T)) > w \\
&\text{then, by congugacy property} \\
&1 - \overline{p}(A_T(S_T) < a_0B(T)) > w \\
&\text{by imprecise probability inequality} \\
&\underline{p}(A_T(S_T) < a_0B(T)) < \overline{p}(A_T(S_T) < a_0B(T)) < 1 - w < w
\end{aligned}
$$

Therefore, only one action could be taken in the imprecise probability asset trading route.

A NPI believer, who prefer to use imprecise expectation operator $\underline{E}$ and $\overline{E}$ could rationally make following decision on asset trading in this scenario at time 0:

Imprecise expectation asset trading route 1.2:

$$
\begin{cases}
\text{Borrow cash } a_0 \text{ and buy the asset } a_0 \text{ at time 0. And sell the asset at time } T \\
\text{return } a_0B(T) \text{ to the lender. if } \underline{E}(A_T) > a_0B(T) \\
\\
\text{Short the asset at for } a_0 \text{ current time, invest the cash } a_0 \text{ at risk free rate and} \\
\text{close the both positions at time } T \text{ if } \overline{E}(A_T) < a_0B(T) \\
\\
\text{No action if none of above satisfied}
\end{cases}
$$

Motivation behind NPI imprecise expectation asset trading route 1.2:

If the lower expectation of asset price A_T at future time T is greater than the value $a_0B(T)$ received when investing a_0 at risk free rate r at time 0, then the profit generated from asset through duration T is expectationally greater than the interest generated from borrowing a_0 at time 0 with risk free rate r. Thus one could rationally borrow cash a_0 and buy the asset a_0 at time 0, expecting profit $\underline{E}(A_T) - a_0B(T)$ at time T. By the similar logic, if $\overline{E}(A_T) < a_0B(T)$, one would rationally short the asset and invest the cash a_0 in risk free rate, expecting profit $a_0B(T) - \underline{E}(A_T)$ at time T.

It is easy to show that only one of actions could be taken in route 1.2: In the imprecise probability framework, one has $\underline{E}(A) \leq \overline{E}(A)$ and therefore $\underline{E}(A_T) > a_0B(T)$ and $\overline{E}(A_T) < a_0B(T)$ could not be satisfied at the same time.

4.2 Simulation of NPI asset trading routes

In this section, we use simulation to study the performance of two proposed NPI trading routes for the asset in the above scenario setting. The goals of the simulation are three-fold. First, to verify the predictive property of NPI imprecise probability in asset trading. Second, to evaluate and compare the performance of different NPI trading routes in asset trading. Third, to identify the effectiveness and efficiency of data learning in NPI imprecise probability.

We only present simulation results with following predefined parameters valued for r,u,d and a_0. Other value of predefined parameters value are also simulated, they all have the similar pattern.

[Predefined parameters value for r,u,d and a_0]:

As stated in the assumptions, we consider the asset over short period of time so the each time step within the short period of time is small. Thus, the discounting rate r is set at $r = 0.0007$ in the simulation. Other parameters value is set as followed, upward movement $u = 1.03$, downward movement $d = 1/u$, initial asset price $a_0 = 100$.

All the trading routes are simulated 100,000 times using the statistical software R version 3.5.1.

4.2.1 Data generation process

Precise data generation process with average market condition

For each simulation trial, the precise data are generated from the family of Bernoulli distribution with random parameter p. To achieve this, in each simulation trial, one firstly draws a random number p from Uniform(0.1,0.9) and then generate $n+T$ data points from Bernoulli(p).

Precise data generation process with specific market condition

For all simulation trials, precise data are generated from one Bernoulli distribution. One sets a number $p \in (0.1, 0.9)$ and then use this predefined p to generate $n+T$ data points from Bernoulli(p) for all simulation trials.

Imprecise data generation process for average market condition

For each simulation trial, imprecise data are generated from the family of Bernoulli distribution with random parameter p_1. One draws a number $p_1 \in$ Uniform$(0.1, 0.9)$ representing the market condition and sets another number $p_2 \in (0, 1)$ representing the "noise" in observation. One then generate 2 arrays $n+T$ data points from Bernoulli(p_1) and Bernoulli(p_2) respectively.

Define ":=" as imprecise data converter, $\binom{1}{0} := \{1\}$, $\binom{0}{0} := \{0\}$, $\binom{1}{1} := \{0, 1\}$ and $\binom{0}{1} := \{0, 1\}$

Imprecise data generation process for specific market condition

For each simulation trial, imprecise data are generated from one Bernoulli distribution. One sets a number $p_1 \in (0.1, 0.9)$ representing the market condition and sets another number $p_2 \in (0, 1)$ representing the "noise" in observation. One then generates 2 arrays of $n+T$ data points from Bernoulli(p_1) and Bernoulli(p_2) respectively.

Define ":=" as imprecise data converter, $\binom{1}{0} := \{1\}$, $\binom{0}{0} := \{0\}$, $\binom{1}{1} := \{0, 1\}$ and $\binom{0}{1} := \{0, 1\}$

4.2.2 Performance evaluation function f_i^A

The performances of NPI asset trading routes are measured by five statistics of the present value pay-off function $f_i^A(n, T, i)$ in 100000 simulations. For each simulation trial $f_i^A(n,T,i)$ is defined as follow:

$$f_i^A(n,T,i) = \begin{cases} A_T(s_T^i)B(T)^{-1} - a_0 & \text{if choose to borrow cash and buy the asset} \\ a_0 - A_T(s_T^i)B(T)^{-1} & \text{if choose to short sell and invest in risk free rate} \\ 0 \quad \text{if no action} \end{cases}$$

where the inputs:
n is the length of historical asset price data one could learn;
T is the future time that the this function is evaluate;
$i \in (1, 100000)$ is the index of that particular simulation trial.

Five performance statistics of this function measure from 100000 simulations are:

Average present value payoff $\overline{f_i^A} = \dfrac{\sum_i f_i^A}{100000}$ Win-loss ratio $R_{wl}^A = \dfrac{|\{i : f_i^A > 0\}|}{|\{i : f_i^A < 0\}|}$

Win rate $R_{wr}^A = \dfrac{|\{i : f_i^A > 0\}|}{100,000}$ Loss rate $R_{lr}^A = \dfrac{|\{i : f_i^A) < 0\}|}{100,000}$

Inaction rate $R_{ir}^A = \dfrac{|\{i : f_i^A = 0\}|}{100,000}$

One should know the sum of win rate and loss rate is not equal to 1, as the NPI trading routes allow "inaction" when the all the desirable events are substantially uncertain.

4.2.3 Sample simulation trials of different asset trading routes given precise or imprecise data

In this subsection, some simple simulation trials are provided to illustrate how each trading route work in the simulation process.

Simulation trials 1 Underlying market condition $p = 0.7$, (For the investor, this information is hidden.), one observes following precise data of a asset in past 7 time stages.

Time stage	-7	-6	-5	-4	-3	-2	-1
Data ($p = 0.7$)	0	1	1	1	0	1	1
Equivalently $(n, j) = (7, 5)$							

One needs to make a decision whether or not enter a risk position of this asset for next 7 time stages. If one is using route 1.1 (imprecise probability trading route) and set threshold value $w = 0.6$. Using the predefined parameters value, one firstly finds out m such that $A_7(m) = a_0 B(7)$. $m \approx 4.3289$, One then find out $m_1 = \lceil m \rceil = 5$, $m_2 = \lfloor m \rfloor = 4$ and calculate $\underline{p}_{(7,5)}(S_7 \geq m_1 = 5) = 0.5 < w$ and $\underline{p}_{(7,5)}(S_7 \leq m_2 = 4) = 0.2797 < w$ so one will take no action in this case.

If one is using route 1.2 (imprecise expectation trading route), using the predefined parameters value, one finds out $\underline{E}_{(7,5)}(A_7) = 105.8076 > a_0 B(7) = 105.022$ and $\overline{E}_{(7,5)}(A_7) = 111.3152 > a_0 B(7) = 105.022$. Thus, one will borrow cash a_0 and buy the asset a_0 at time 0 and close the risk position at time 7.

Simulation trial 2 Underlying market condition $p = 0.3$, one observes following data of a asset in past 7 time stages

Time stage	-7	-6	-5	-4	-3	-2	-1
Data ($p = 0.3$)	0	0	1	0	0	0	1
Equivalently $(n, j) = (7, 2)$							

One needs to decide whether or not enter a risk position of this asset for next 7 time stages. If one is using route 1.1 (imprecise probability trading route) and set threshold value $w = 0.6$. Using the predefined parameters value, one firstly finds out m such that $A_7(m) = a_0 B(7)$. $m \approx 4.3289$. One then find out $m_1 = \lceil m \rceil = 5$, $m_2 = \lfloor m \rfloor = 4$ and calculate $\underline{p}_{(7,2)}(S_7 \geq m_1 = 5) = 0.0513 < w$ and $\underline{p}_{(7,2)}(S_7 \leq m_2 = 4) = 0.8569 > w$. Thus one will Short the asset at for a_0 current time, invest the cash a_0 at risk free rate and close the both positions at time 7 in this case.

If one is using route 1.2 (imprecise expectation trading route), using the predefined parameters value, one firstly finds out $\underline{E}_{(7,2)}(A_7) = 90.52443 < a_0 B(7) = 105.022$ $\overline{E}_{(7,2)}(A_7) = 95.41828 < a_0 B(7) = 105.022$. Thus, one will short the asset for a_0 at time 0, invest the cash a_0 at risk free rate and close the both positions at time 7

Simulation trial 3 Underlying market condition $p_1 = 0.7$, and noise level $p_2 = 0.2$, one observes following data of a asset in past 7 time stages.

Time stage	-7	-6	-5	-4	-3	-2	-1
Data ($p_1 = 0.7$, $p_2 = 0.2$)	0	1	1	1	{0,1}	1	1
Equivalently $[n, J] = [7, (5, 6)]$							

One needs to make a decision whether or not enter a risk position of this asset for next 7 time stages. If one is using route 1.1 (imprecise probability trading route) and set threshold value $w = 0.6$. Using the predefined parameters value, one firstly finds out m such that $A_7(m) = a_0B(7)$. $m \approx 4.3289$. One then find out $m_1 = \lceil m \rceil = 5$, $m_2 = \lfloor m \rfloor = 4$ and calculate $\underline{p}_{[7,(5,6)]}(S_7 \geq m_1 = 5) = 0.5 < w$ and $\underline{p}_{[7,(5,6)]}(S_7 \leq m_2 = 4) = 0.0962 < w$. Therefore one will take no action in this case.

If one is using route 1.2 (imprecise expectation trading route), using the predefined parameters value, one firstly finds out $\underline{E}_{[7,(5,6)]}(A_7) = 105.8076 > a_0B(7) = 105.022$ and $\overline{E}_{[7,(5,6)]}(A_7) = 117.0397 > a_0B(7) = 105.022$. Thus, one will borrow cash a_0 and buy the asset a_0 at time 0 and close the risk position at time 7.

Simulation trial 4 Underlying market condition $p_1 = 0.3$, and noise level $p_2 = 0.6$, one observes following data of a asset in past 7 time stages.

Time stage	-7	-6	-5	-4	-3	-2	-1
Data ($p_1 = 0.7$, $p_2 = 0.6$)	0	1	0	1	{0,1}	{0,1}	{0,1}
Equivalently $[n, J] = [7, (2, 5)]$							

One needs to make a decision whether or not enter a risk position of this asset for next 7 time stages. If one is using route 1.1 (imprecise probability route) and set threshold value $w = 0.6$. Using the predefined parameters value, one firstly finds out m such that $A_7(m) = a_0B(7)$. $m \approx 4.3289$. One then find out $m_1 = \lceil m \rceil = 5$, $m_2 = \lfloor m \rfloor = 4$ and calculate $\underline{p}_{[7,(2,5)]}(S_7 \geq m_1 = 5) = 0.0513 < w$ and $\underline{p}_{[7,(2,5)]}(S_7 \leq m_2 = 4) = 0.2797 < w$. Consequently, one will take no action in this case.

If one is using route 1.2 (imprecise probability route), using the predefined parameters value, one firstly finds out $\underline{E}_{[7,(2,5)]}(A_7) = 90.52443 < a_0B(7) = 105.022$ $\overline{E}_{[7,(2,5)]}(A_7) = 111.3152 > a_0B(7) = 105.022$. Thus, one will take no action in this case.

4.2.4 Performance of NPI asset trading routes under average market condition given precise data available

In this section, under average market condition, given precise data available, the performance of NPI asset trading routes are evaluated and discussed.

Figure 4.1 shows that for different combinations of precise historical data point n and future length T, both routes 1.1 and 1.2 yield positive average present value payoff on average market condition. It could be noticed that, with small amounts of historical data n available, route 1.2 performs better in term of average present value payoff. The reason for this is that route 1.1 is a more conservative trading route which tends to avoid making trading when information learnt from data is not very sufficient. This is further confirmed in Figure 4.2.

Figure 4.2 demonstrates the performance of decision routes 1.1 and 1.2 in terms of five performance index $\overline{f_i^A}$, R_{wl}^A, R_{wr}^A, R_{lr}^A and R_{ir}^A at time $T = 100$ under average market condition.

It could be observed from all indexes that both NPI based asset trading routes have very quick learning speed. The average present value payoff $\overline{f_i^A}$ from both routes increases sharply when a small number of data become available. Moreover, all the performance indexes for both trading routes become better when more data is presented and they stabilized after 20 data point are available. Overall, in terms of long run payoff $\overline{f_i^A}$, route 1.2 slightly outperforms route 1.1. However, as previously mentioned, route 1.1 is a more conservative trading route which tends to avoid making trading when information learned from data is not very sufficient. This indeed is the case, as one could observe from R_{ir}^A plot in Figure 4.2. Route 1.1 overall has higher inaction rate then route 1.2, especially in the case only small amount of data is present. Also lower loss rate R_{lr}^A in route 1.1 could be observed from Figure 4.2. This is as expected since route 1.2 prioritize the expectation of

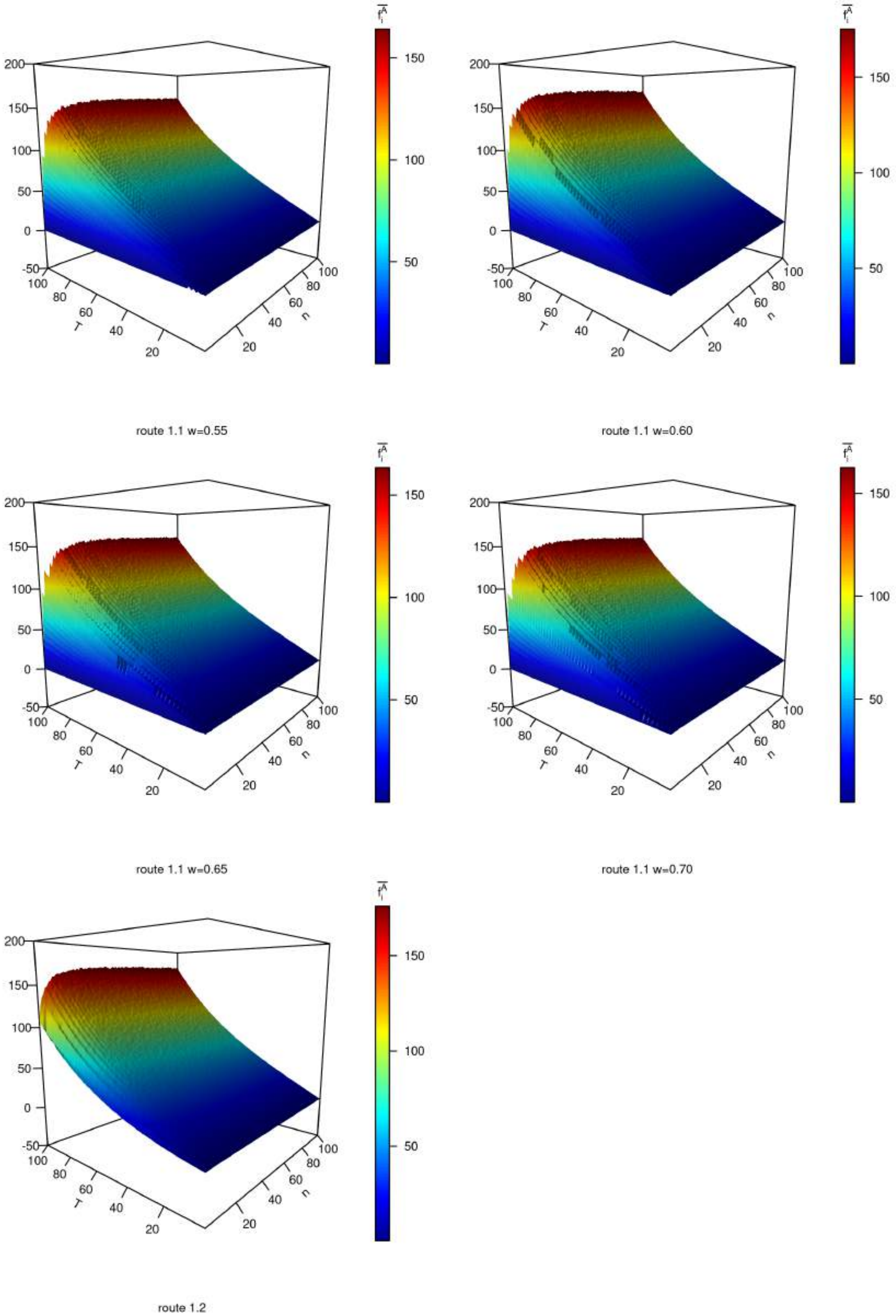

Figure 4.1: APVP of routes 1.1 and 1.2 under average market condition given precise data (APVP stands for average present value payoff in all following figures).

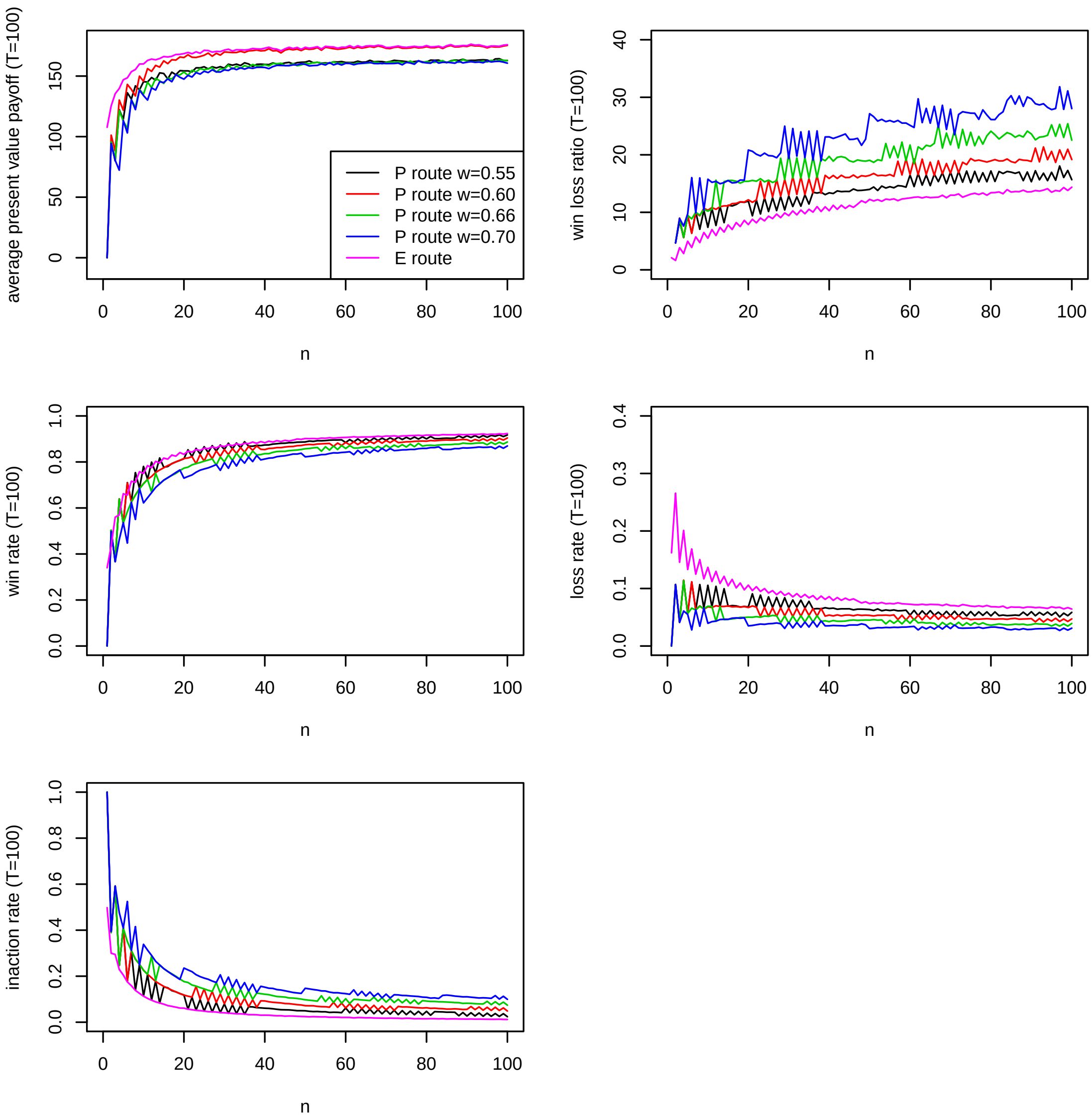

Figure 4.2: Performance comparison of routes 1.1 and 1.2 at $T = 100$ under average market condition given precise data.

profit while route 1.1 prioritize achieve profit with a certain probability threshold. With adjustment of threshold parameter w, route 1.1 is able to have better control in win loss ratio R^A_{wl} and loss ratio R^A_{lr} than route 1.2. At threshold parameter $w = 0.7$, route 1.1 could achieve loss rate less than 0.1 and win loss ratio greater than 20 when $n \geq 50$ data points are available. Although route 1.1 has lower win rate generally, one should attribute this to its higher inaction rate R^A_{lr} as it try to avoid taking action when the event $A_T(S_T) > a_0B(T)$ or $A_T(S_T) < a_0B(T)$ is not certain up to the threshold value w. Albeit route 1.1 have better performance in win loss ratio R^A_{wl} and loss ratio R^A_{lr}, one should not neglect that performance of route 1.2 in those indexes are still very satisfactory.

Overall, given precise data, under average market condition, both trading routes yield positive average present value payoff and have good performance in R^A_{wr}, R^A_{wl} and R^A_{lr}. Both trading routes are able to effectively learn information from the historical data and execute correct action accordingly. When more data become available, the performances of both trading routes become better. By avoiding taking action when the desired event is not certain up to a level, route 1.1 provide good control on the loss rate and win-loss ratio when small amount of data is available. In contrast, route 1.2 has higher present value payoff in long run with less attention in the risk control.

4.2.5 Performance of NPI asset trading routes under different market conditions given precise data available

The previous section showed that given precise data, NPI asset trading routes are well performed under average market condition. This section investigates the next level of detail by evaluating the performance under a specific market condition.

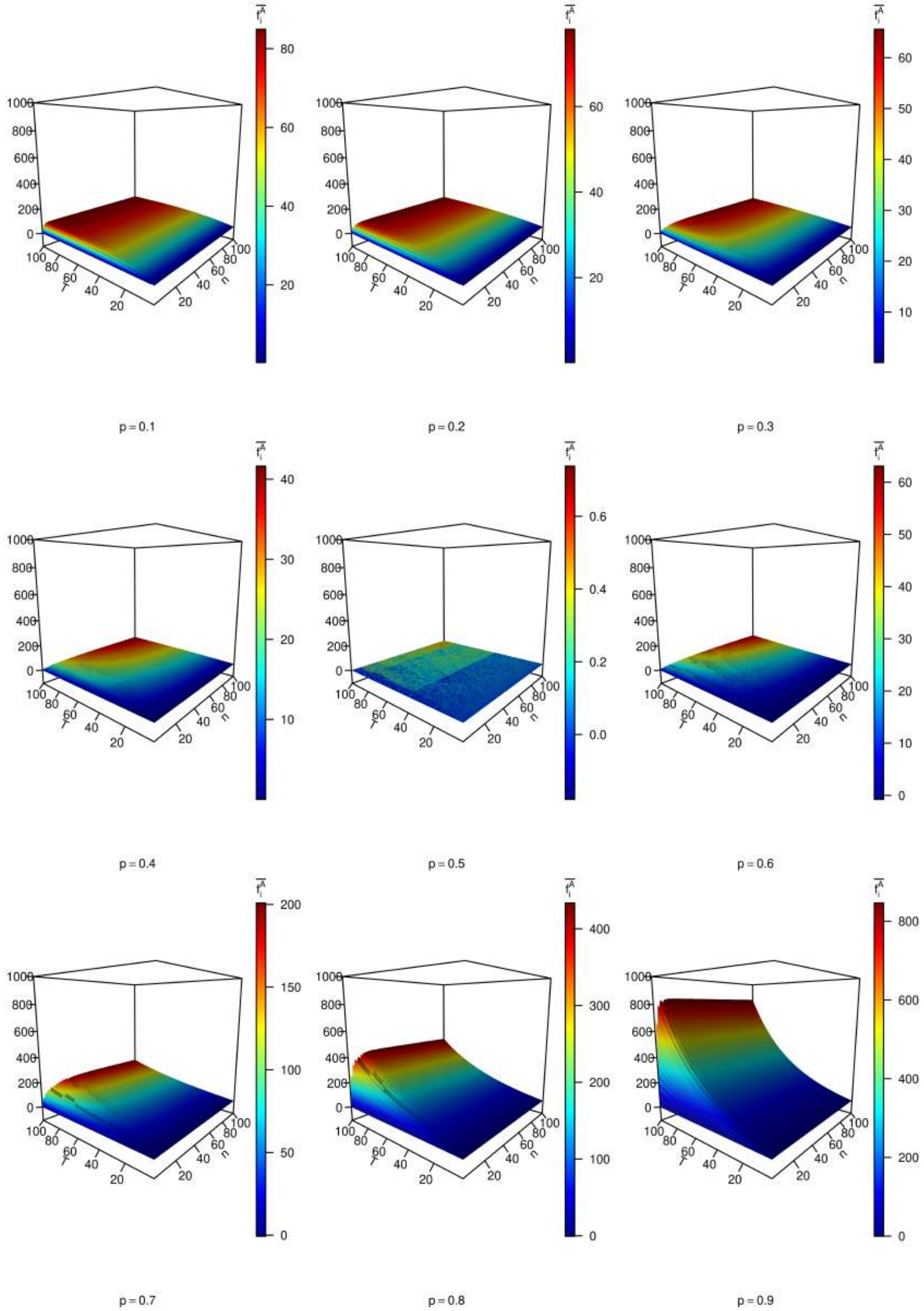

Figure 4.3: Under different market condition, APVP of trading route 1.1 with threshold value $w = 0.6$.

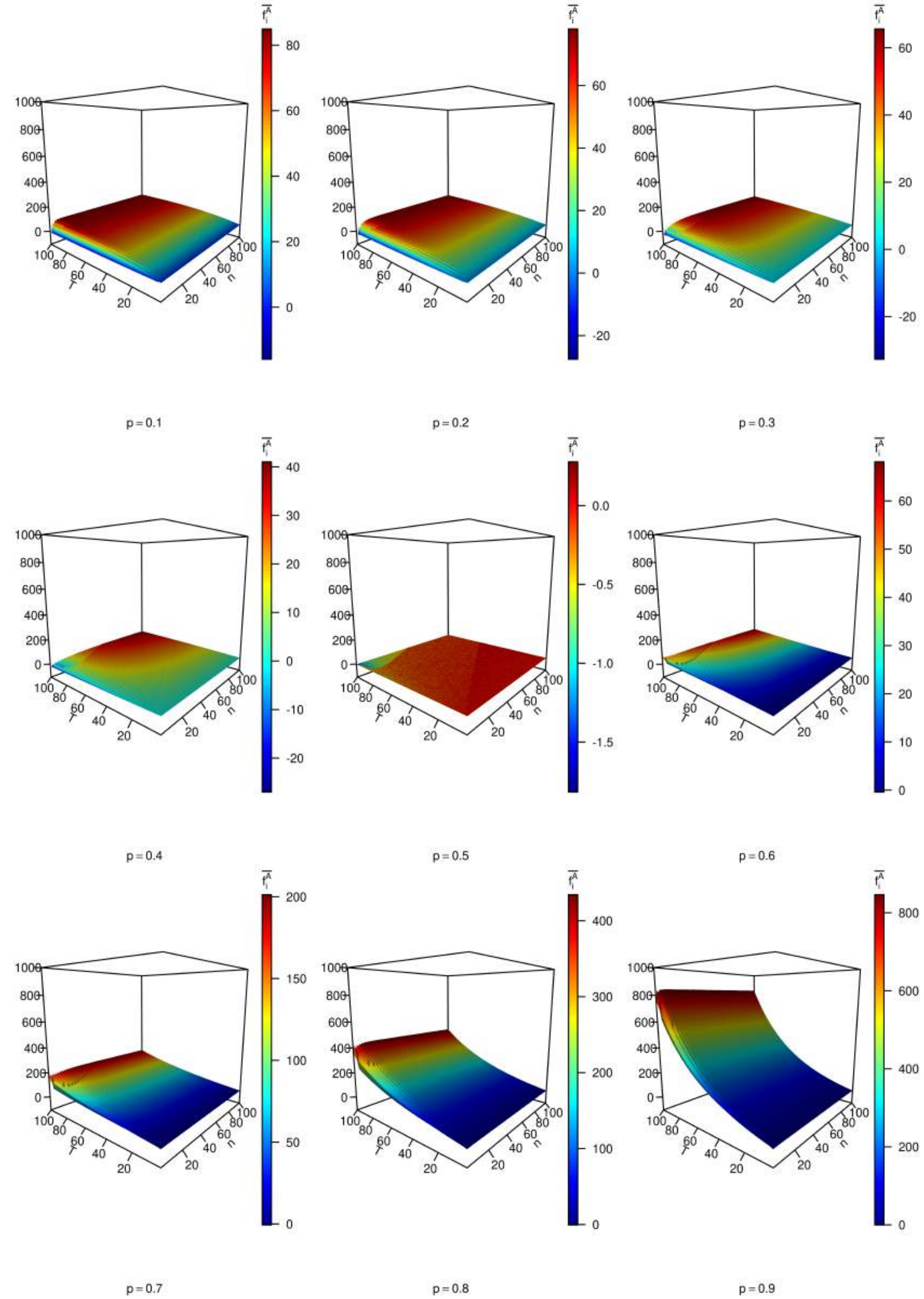

Figure 4.4: Under different market condition, APVP of trading route 1.2.

From Figure 4.3 and 4.4, it could be concluded that both trading routes readily recognize the market condition when a few data become available and have good performance when the market is relatively one-sided $p \in (0.1, 0.4)$ or $p \in (0.6.0.9)$. Especially one should notice that route 1.1 preserves the positivity of average present value payoff throughout all different market condition due to its risk control nature. On the other hands, route 1.2 appears to has negative average present value payoff under market condition $p \in (0.2, 0.5)$ when the number of data is insufficient. Nevertheless, it is able to rectify its trading strategy when more data become available which results in positivity on the rest of the area of present value payoff surface.

The maximum present value payoff could be achieved by short selling is less than the initial asset price $a_0 = 100$. When the market is in recession $p \in (0.1, 0.4)$, Both trading routes could recognize the market condition from the data and did correct short selling in most of the case. From Figure 4.3 and 4.4, one could observed that with 40 data point available, under market condition $p \in (0.1, 0.4)$, both trading strategy could achieve approximately 20% of average present value payoff after time $T = 80$.

When the market is in upswing $p \in (0.6, 0.9)$, both trading routes are able to recognize the market condition when small of data $n \geq 15$ are available. The second action, borrowing cash and buying the asset, is executed frequently. The best performances of the average present value payoff $\overline{f_i^A}$ for both trading routes occur when the market condition is $p = 0.9$

When the market condition is neutral $p = 0.5$, the correct trading action is taking no action as the market has no obvious trend. Both trading routes are able to execute correct action in most the case under this situation, resulting a flat surface of $\overline{f_i^A}$ in Figure 4.3 and 4.4 when $p = 0.5$.

To see more detail and have a more direct comparison of trading routes' performances, Figure 4.5-4.9 are plotted, which present all the performance indexes at time $T = 100$ under different market conditions for both trading routes.

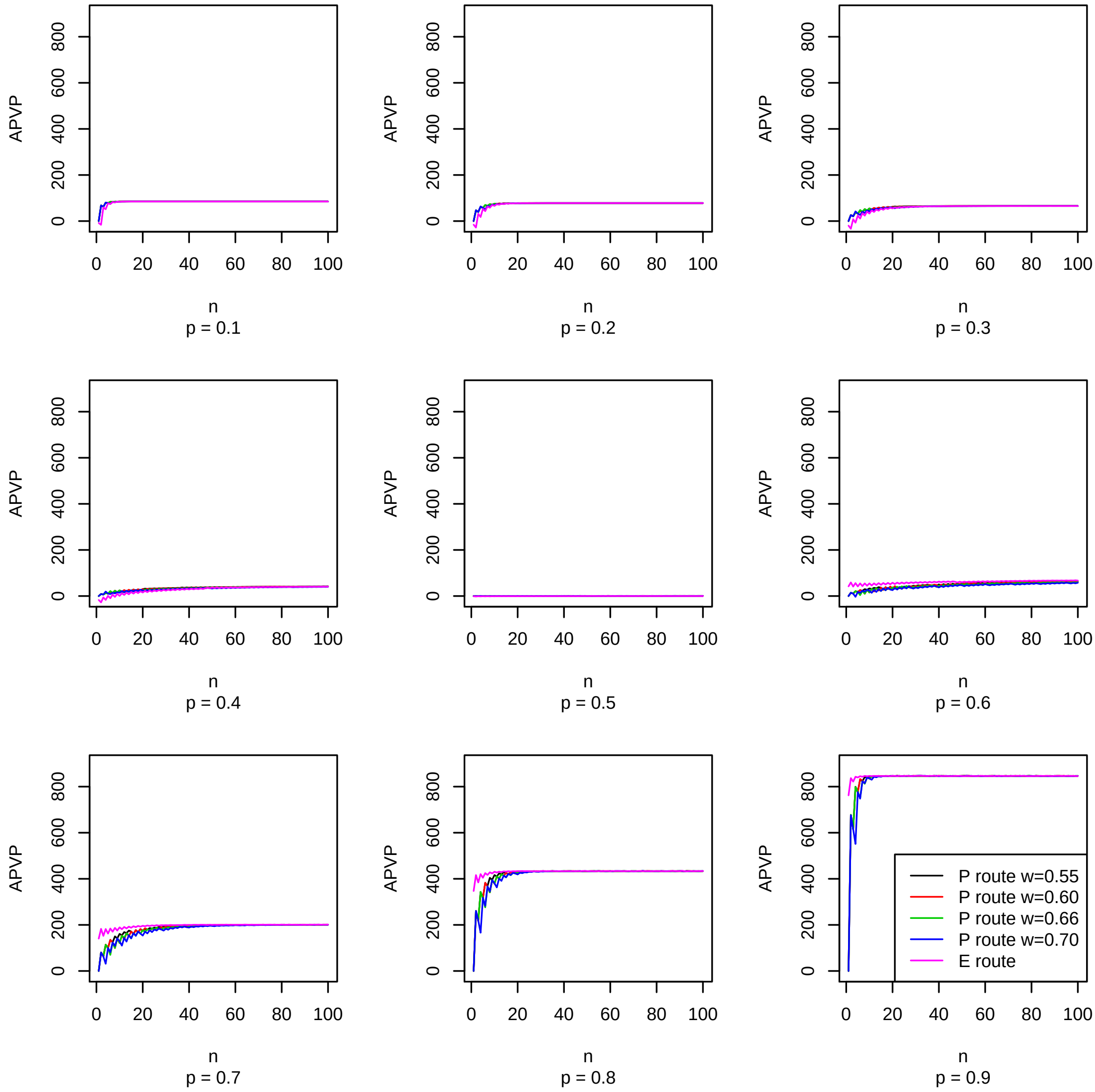

Figure 4.5: Under different market condition, APVP of trading routes 1.1 and 1.2.

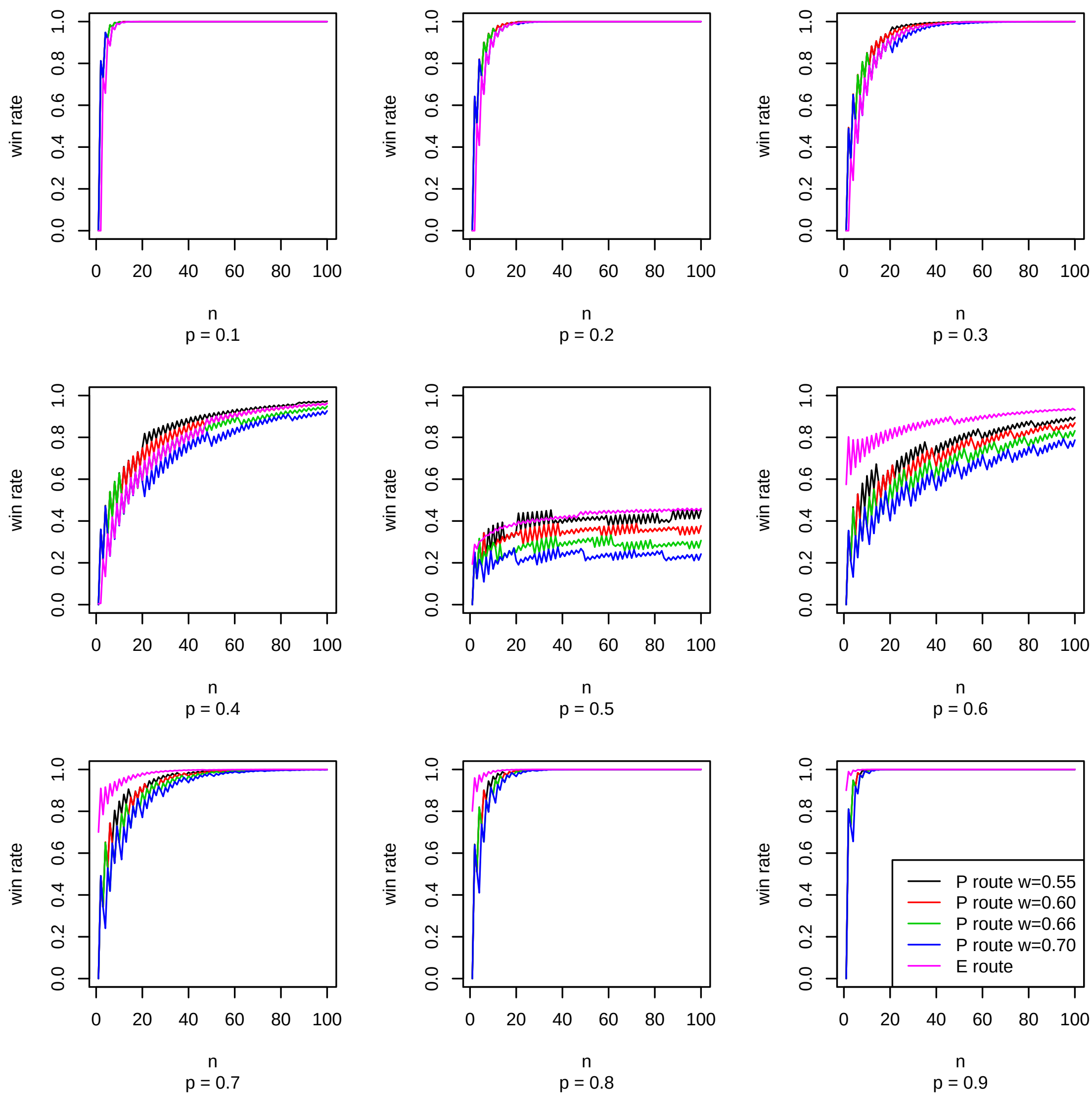

Figure 4.6: Under different market condition, WR of trading routes 1.1 and 1.2 (WR stands for win rate in all following figures).

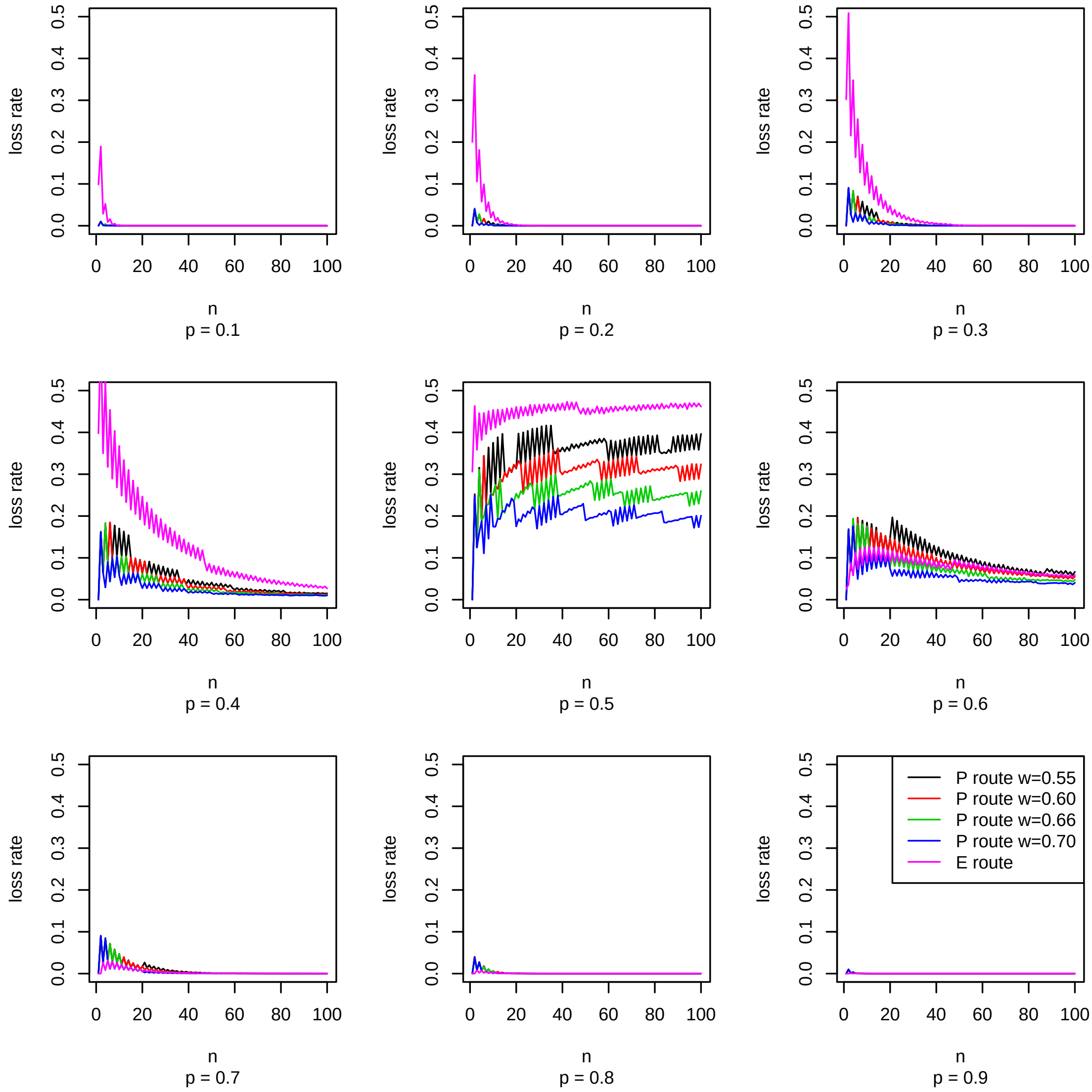

Figure 4.7: Under different market condition, LR of trading routes 1.1 and 1.2 (LR stands for loss rate in all following figures).

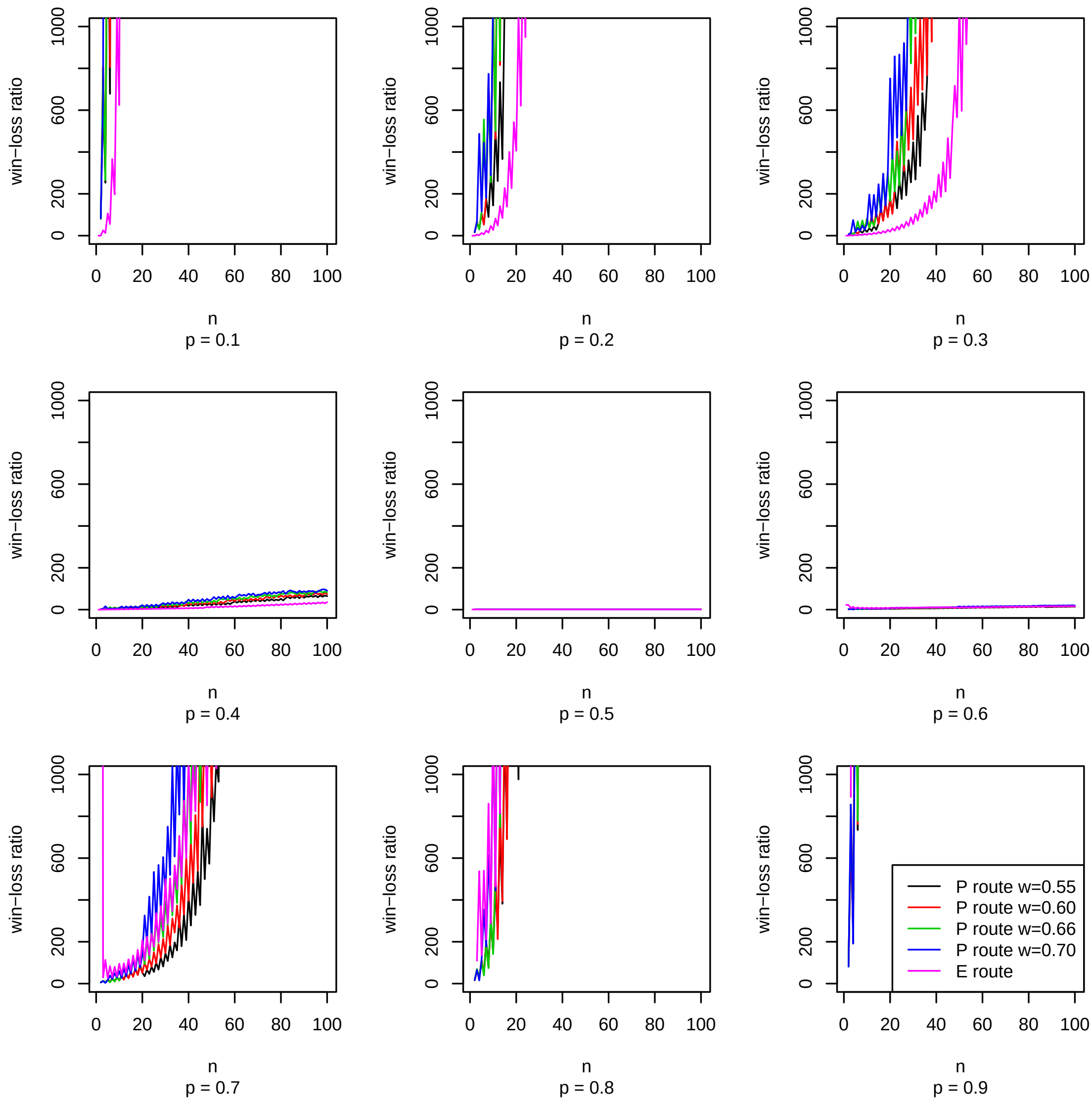

Figure 4.8: Under different market condition, WLR of trading routes 1.1 and 1.2 (WLR stands for win-loss ratio in all following figures).

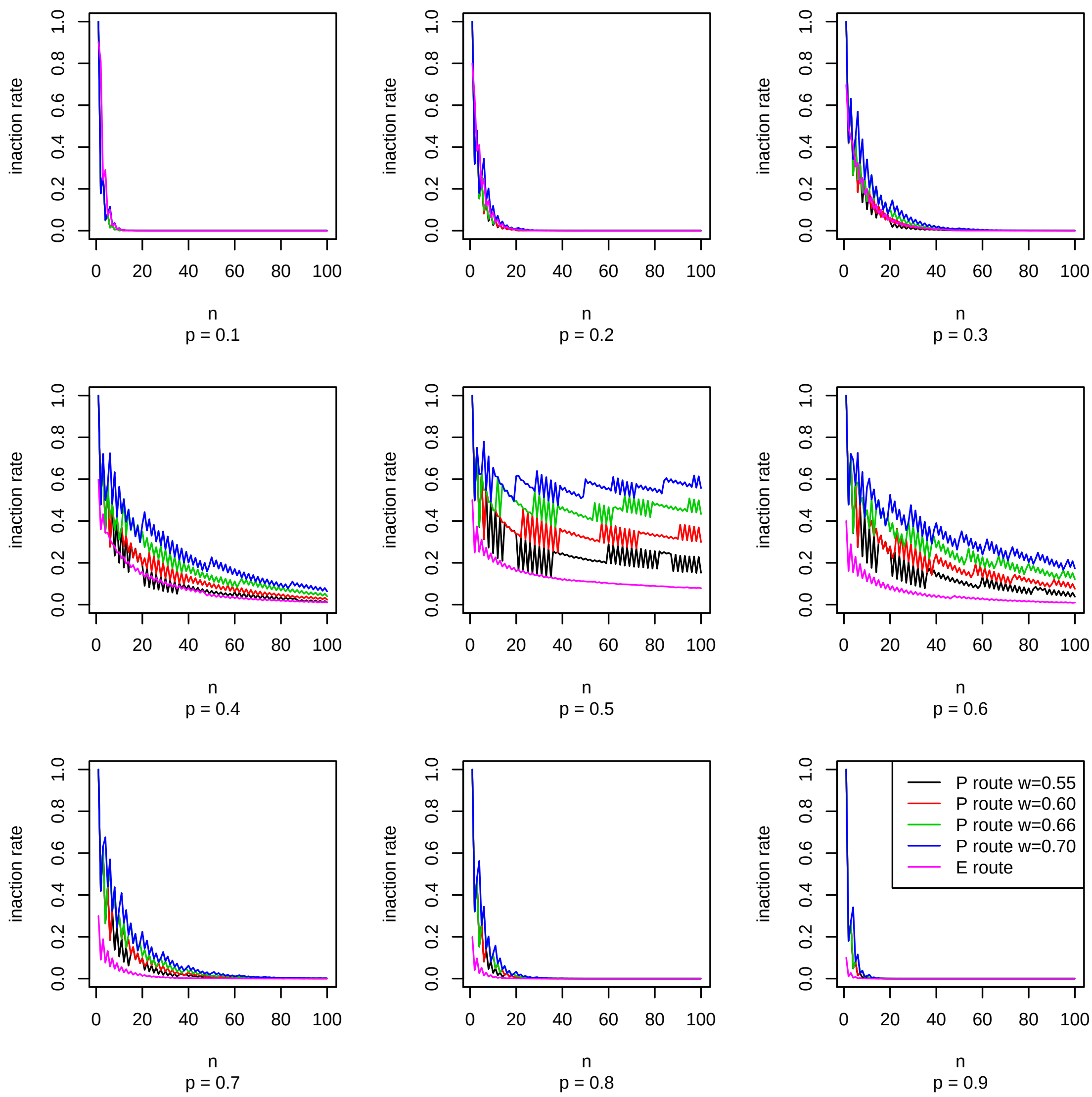

Figure 4.9: Under different market condition, IR of trading routes 1.1 and 1.2 (IR stands for inaction rate in all following figures).

From Figure 4.5, it could be confirmed again both trading routes are able to execute correct action when the market condition is relatively one-sided. As we expected previously, both trading routes avoid taking unreasonable actions when market condition is neutral, indicated by the higher inaction rate when market condition is $p \in (0.4, 0.6)$ in Figure 4.9. Moreover, from Figure 4.5, one could know route 1.2 indeed prioritizes on maximizing the present value payoff in the long run. It generally has higher $\overline{f_i^A}$ than route 1.1 under different market conditions. However, from Figure 4.7, although both trading routes have similar loss rate when market condition is favour for buying the asset, route 1.1 actually has better risk control than route 1.2 in loss rate when the market condition is in favor of short selling the asset or has no trend $p \in (0.1, 0.5)$. Also, with a higher value of threshold value w set in route 1.1, route 1.1 tends to have higher inaction rate and higher win-loss ratio throughout all different market conditions, indicated by Figure 4.8 and Figure 4.9

4.2.6 Performance of NPI asset trading routes under average market condition given imprecise data available

In this section, given imprecise data, under average market condition, the performances of both NPI asset trading routes are evaluated.

From Figure 4.10 and Figure 4.11, it could be observed that under the average market condition, both route 1.1 and route 1.2 preserve positivity on average present value payoff regardless what noise level contains in the data. With noise level $p_2 = 0.1$, the positive average present value payoff surface of both trading routes resemble the corresponding surface in Figure 4.1 which has no noise in data. As the noise level increase, the positive average present value payoff surface from both trading routes are flattened due to more inactions are taken in the trading. This indicates both NPI asset trading routes effectively recognize the noise level in the data and are able to adjust its trading action correspondingly.

Figures 4.12-4.15 give more detail of the performances of both trading routes by plotting $\overline{f_i^A}$, R^A_{wr}, R^A_{lr} and R^A_{ir} at time $T = 100$ with different number of data point n available. Win-loss ratio R^A_{wl} is not presented in this section, because as noise level gradually increases, the inaction rate R^A_{ir} increases to nearly 1 and loss rate R^A_{lr} drop to nearly 0, which eventually makes Win-loss ratio R^A_{wl} “blow up”.

As the noise level increase, the information contained in the data becomes insufficient for one to make a sensible decision and both trading routes heuristically choose to take no action. This can be seen in Figure 4.15. The inaction rate R^A_{ir} of both trading routes increases dramatically as the noise level p_2 increases. At noise level $p_2 = 0.5$, the inaction rates of both trading routes are asymptotic to 1 after 50 data points become available.

With a relatively lower level of noise presented $p_2 \in (0.1, 0.3)$, both trading

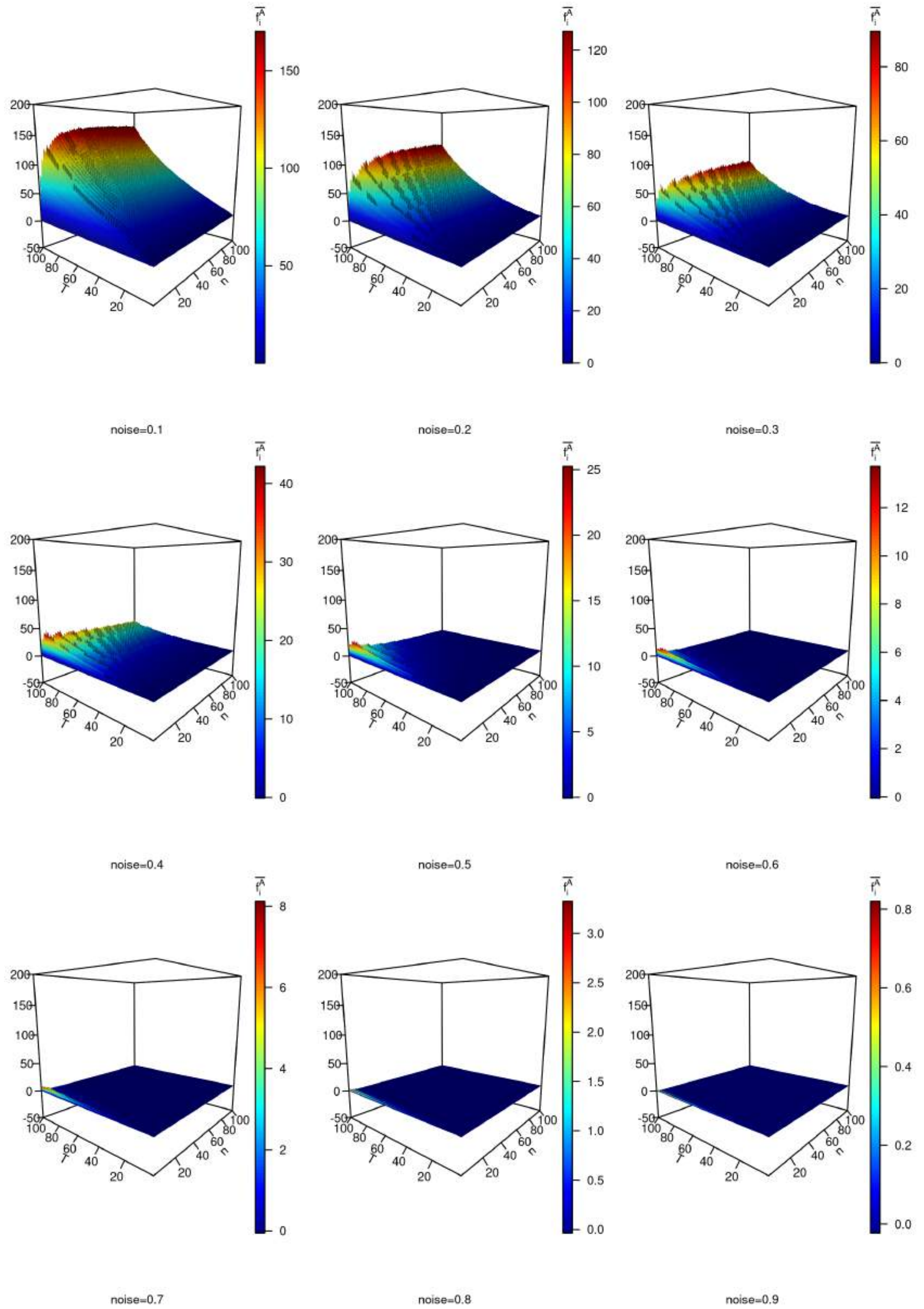

Figure 4.10: With average market condition, APVP of trading route 1.1 with threshold value $w = 0.6$ under different noise levels p_2.

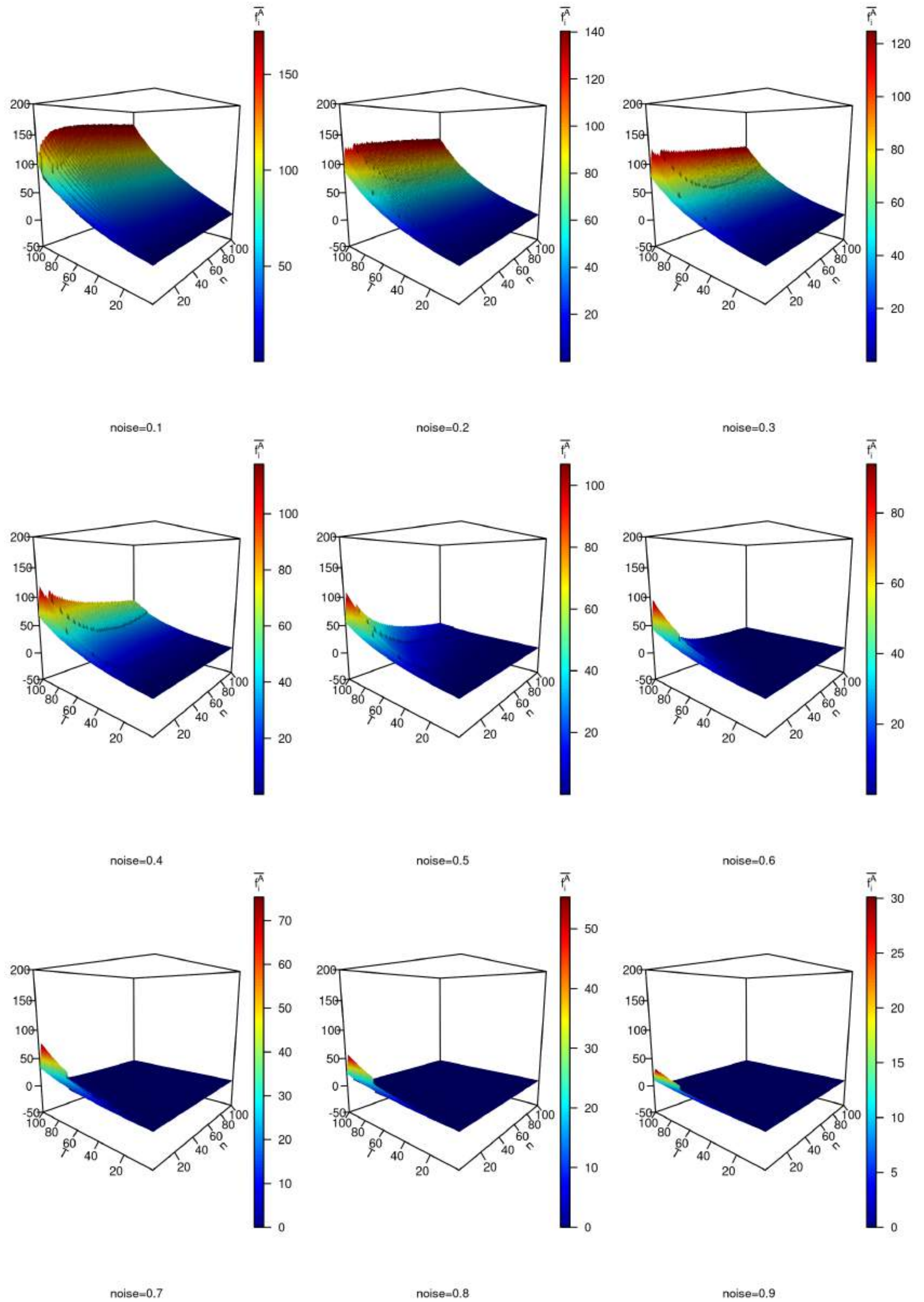

Figure 4.11: With average market condition, APVP of trading route 1.2 under different noise levels p_2.

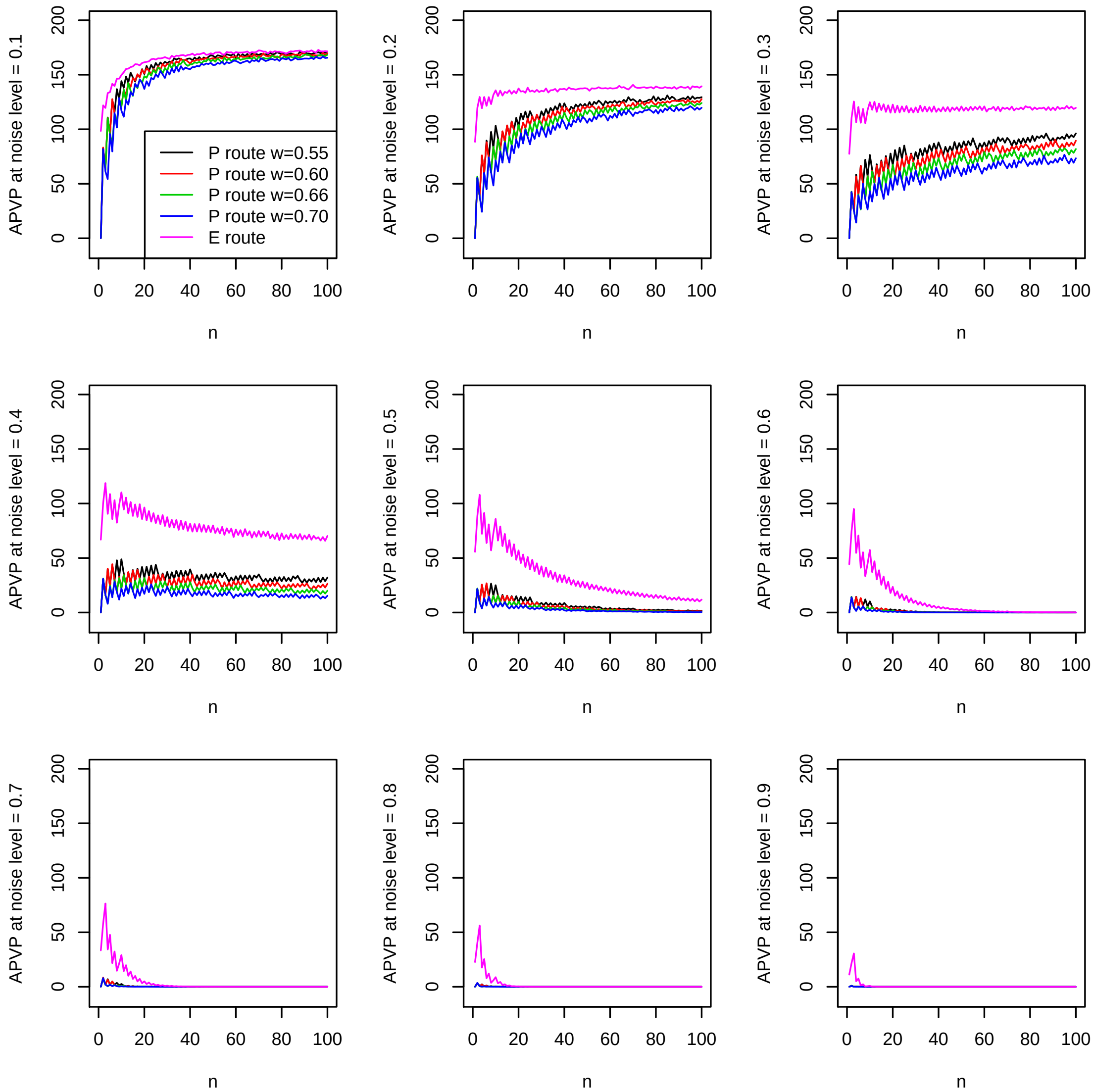

Figure 4.12: With average market condition, at time $T = 100$, APVP of both trading routes 1.1 and 1.2 under different noise levels p_2.

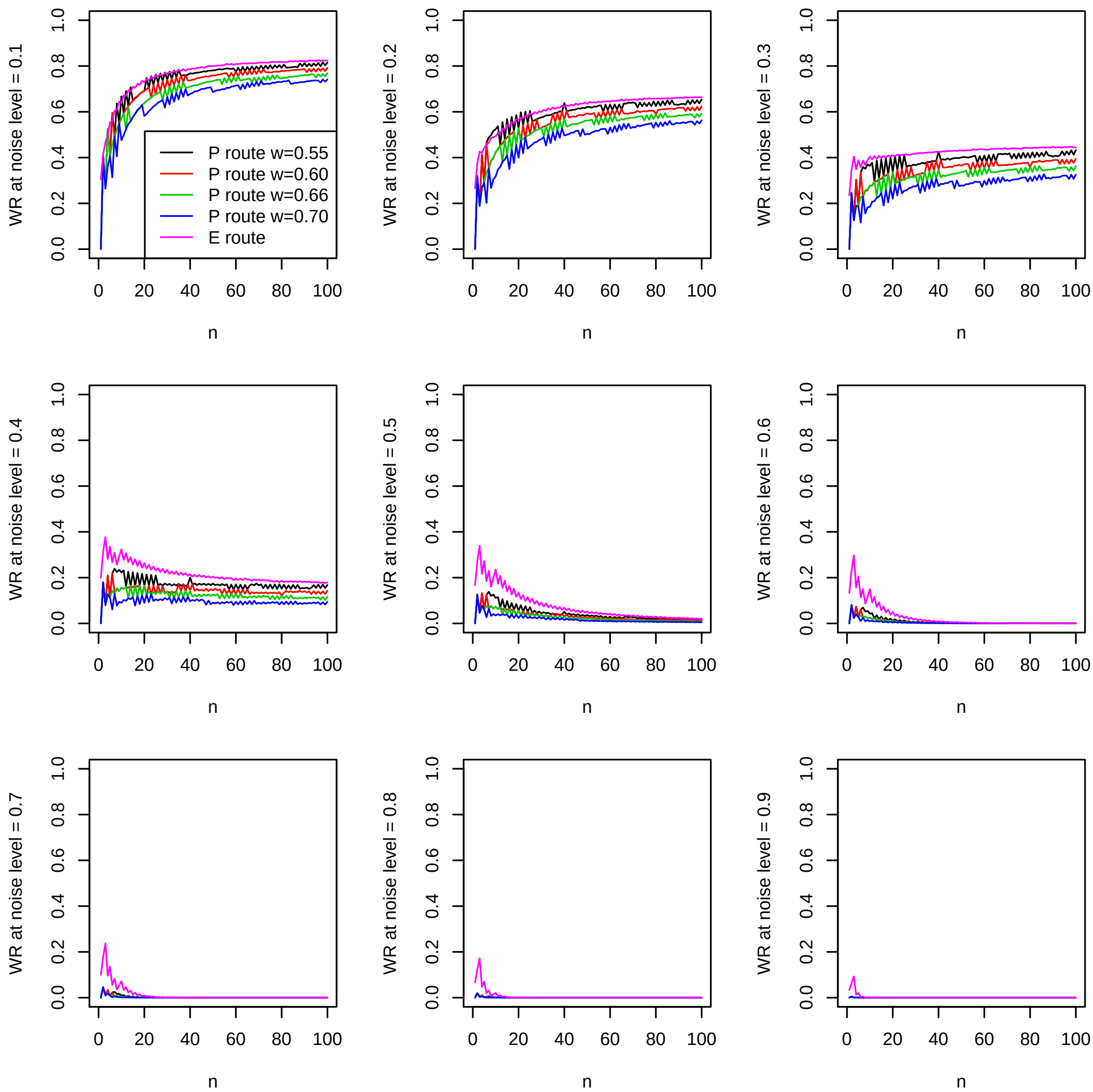

Figure 4.13: With average market condition, at time $T = 100$, WR of both trading routes 1.1 and 1.2 under different noise levels p_2.

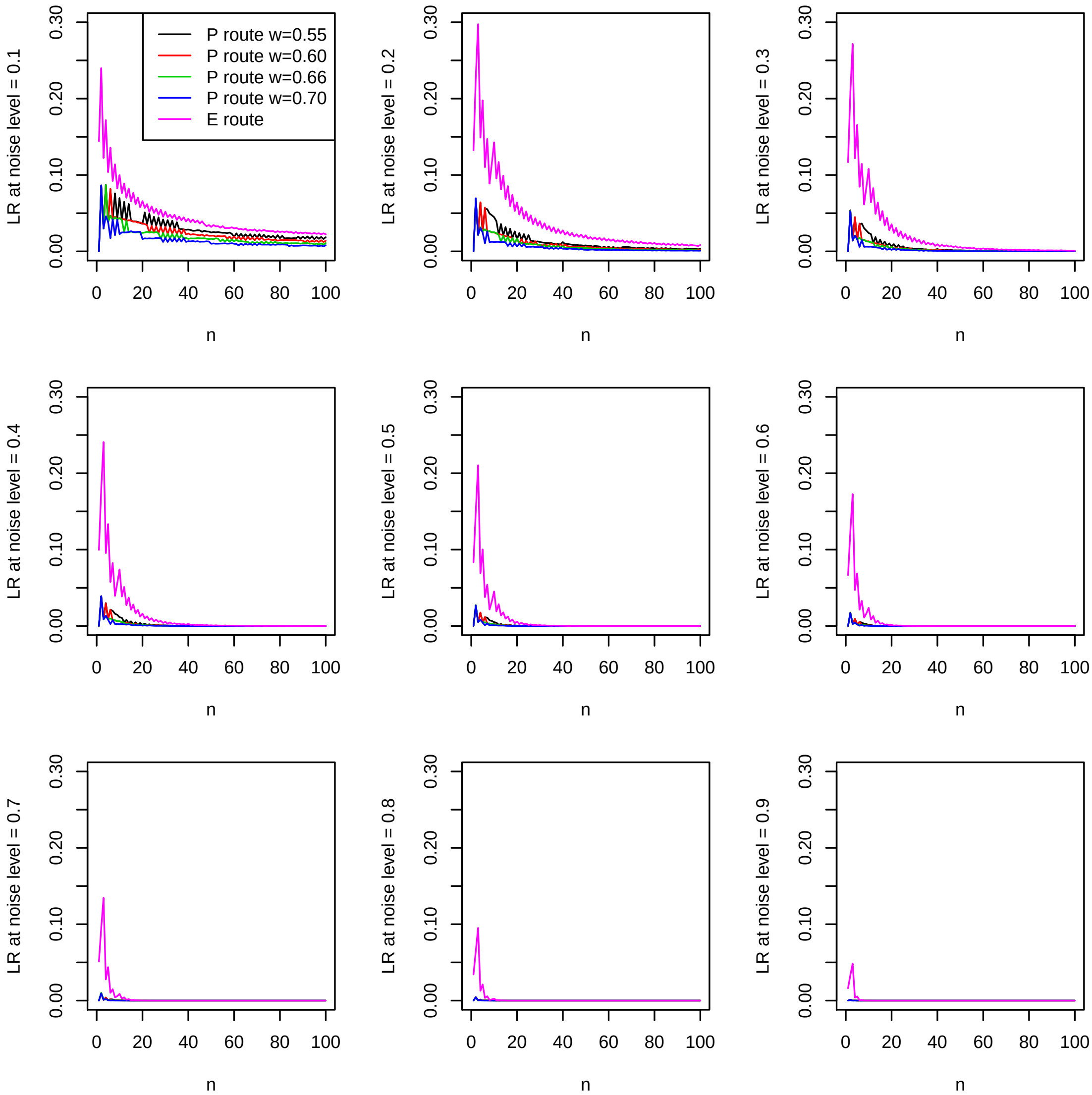

Figure 4.14: With average market condition, at time $T = 100$, LR of both trading routes 1.1 and 1.2 under different noise levels p_2.

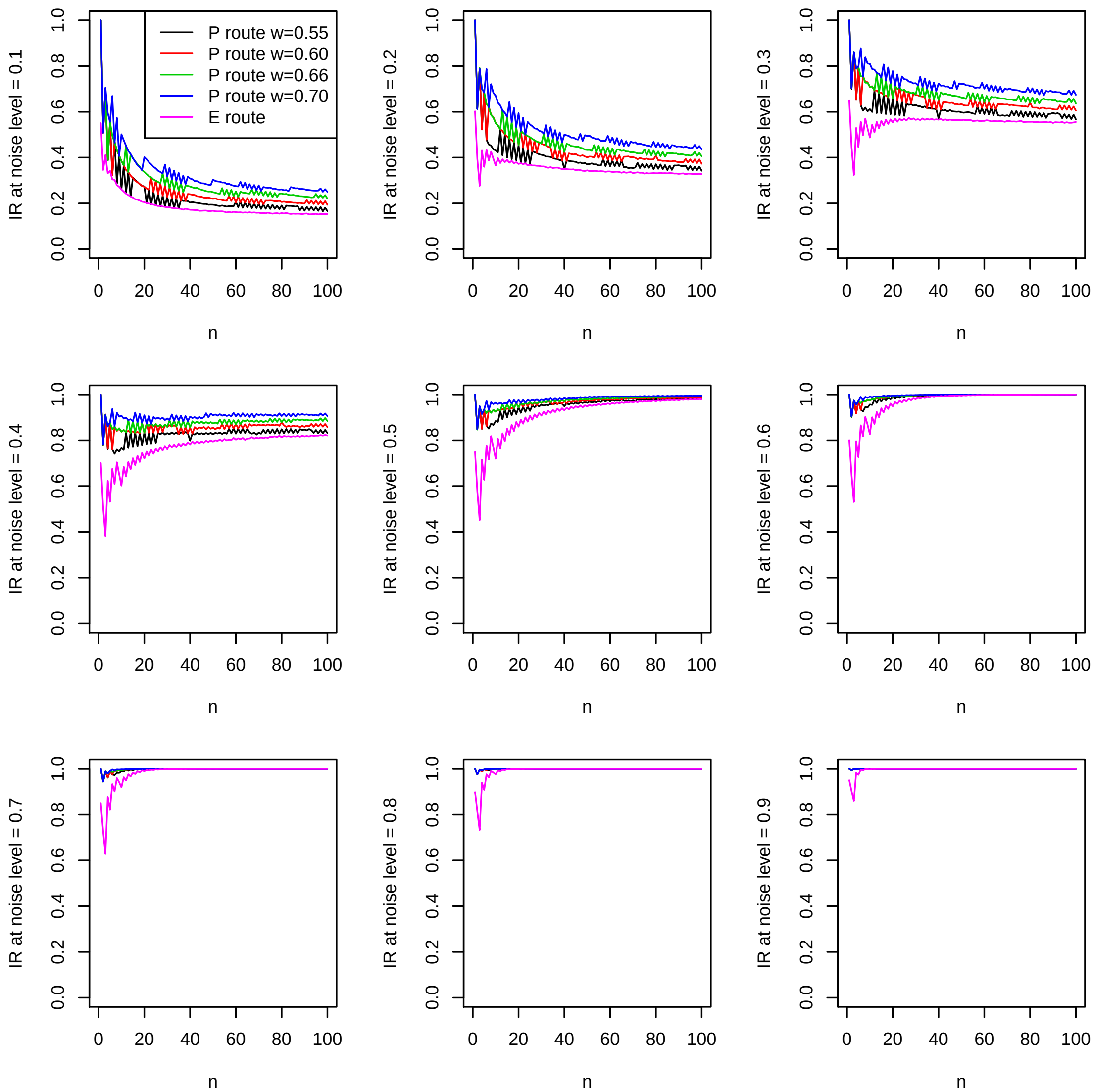

Figure 4.15: With average market condition, at time $T = 100$, IR of both trading routes 1.1 and 1.2 under different noise levels p_2.

routes are still able to extract useful information from the data. Therefore, as the number of low level noise data increase, both trading routes gradually recognize underlying distribution and start to take more actions and the performances of all index resemble the corresponding precise case in Section 4.2.4.

In contrast, when a higher level of noise presented in the data $p_2 \in (0.4, 0.9)$, given sufficient data are available ($n > 20$), both trading routes realized the information contained in the data is too ambiguous and avoid to take action in most cases. (See Figure 4.15)

It could be seen from Figures 4.12-4.14 that both trading routes still maintain their primary objective respectively under low noise affection. Namely, route 1.1 emphasize risk control on loss rate and win-loss ratio while route 1.2 focuses on achieving maximum average present value payoff.

4.2.7 Performance of NPI asset trading routes different market under different market conditions given imprecise data available

In this section, given imprecise data, the performance of NPI asset trading routes is further evaluated different market under different market conditions.

It is observed from simulations that both NPI trading routes, under all market condition $p_1 \in (0.1, 0.9)$, are able to effectively and efficiently recognize the noise from the imprecise data and gradually take less trading action as the noise level increase. When low noise level is presented $p_2 \in (0.1, 0.4)$, both trading routes are able to recognize the underlying market condition and execute correct action accordingly. Since both trading routes share similar patterns of decaying phenomenon on the average present value $\overline{f_i^A}$ surface, and one complete trading route example requires nine pages of space, for the sake of brevity, we only present one complete example of average present value $\overline{f_i^A}$ surface for route 1.1 with threshold value $w = 0.6$ in the Appendix A. (See Figures A.1-A.9)

Both route 1.1 and route 1.2 maintain their respective primary objectives in trading under all market conditions when only low noise level is presented. As a example, the performance indexes of $\overline{f_i^A}$, R_{wr}^A, R_{lr}^A and R_{ir}^A under market condition $p_1 = 0.9$ at time $T = 100$ for both trading routes is demonstrated below. (See Figures 4.16-4.19). The win-loss ratio profile is omitted for the same reason mentioned in previous section.

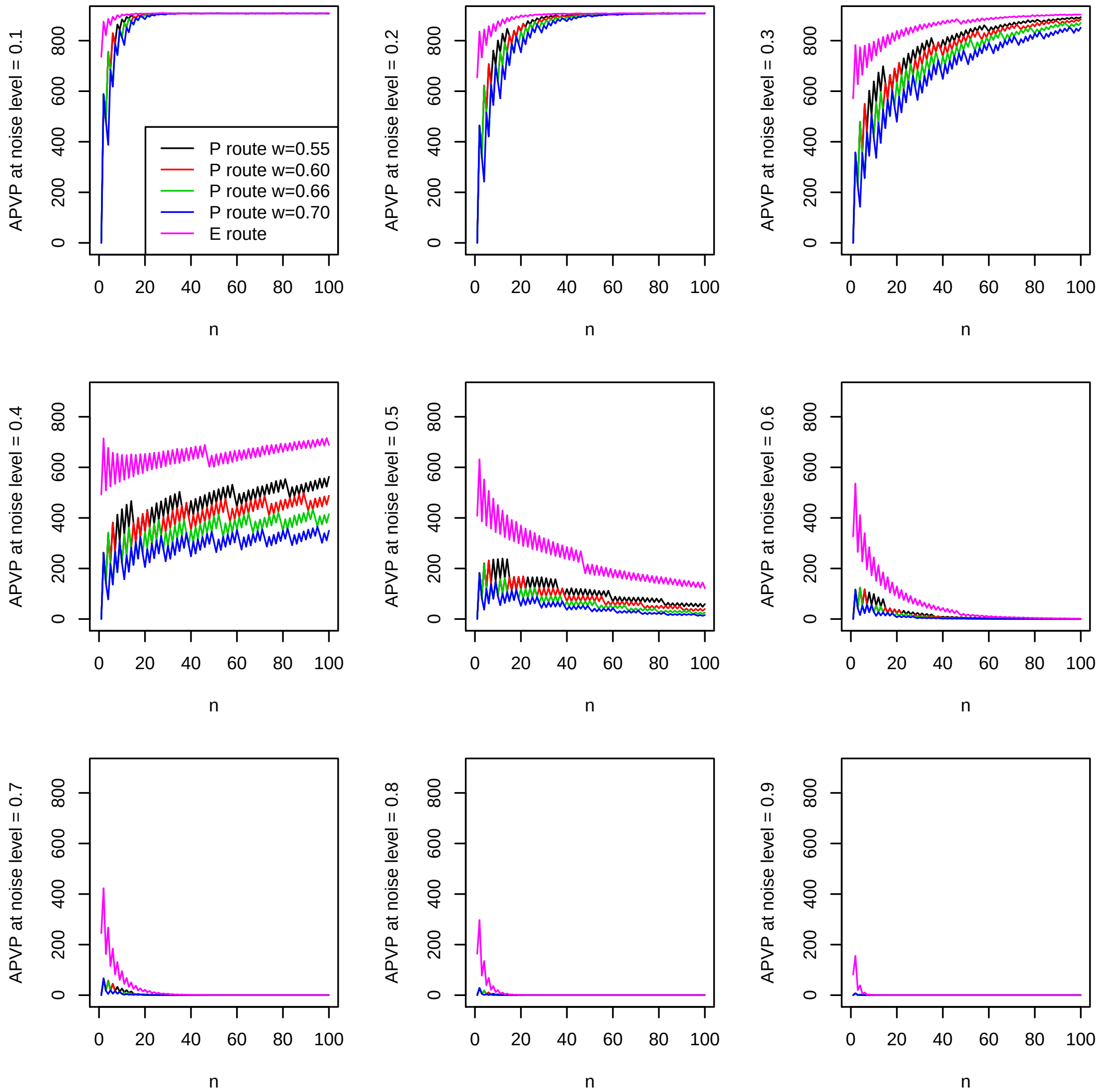

Figure 4.16: APVP of routes 1.1 and 1.2 at $T = 100$ under market condition $p_1 = 0.9$ and different noise levels p_2.

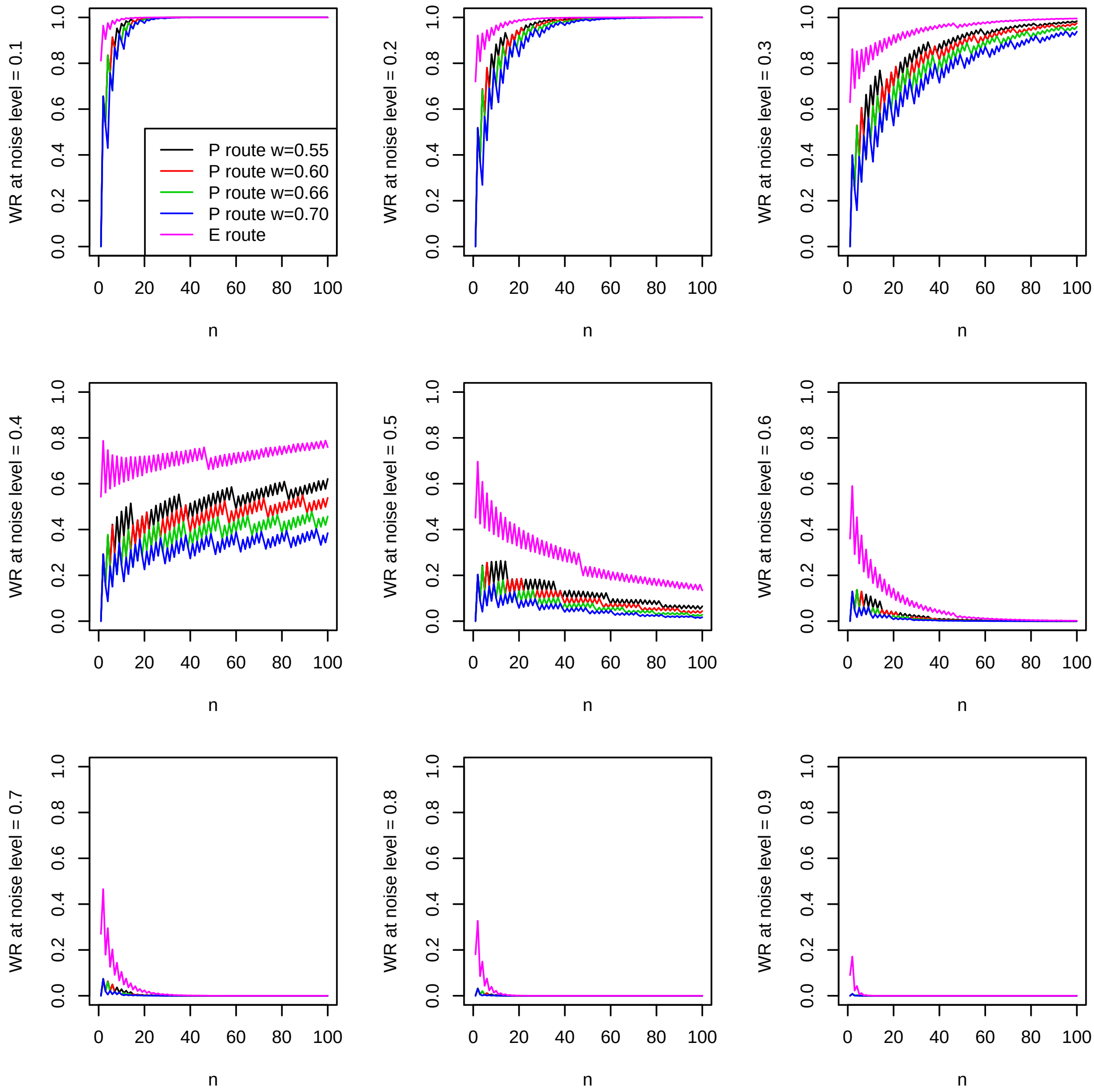

Figure 4.17: WR of routes 1.1 and 1.2 at $T = 100$ under market condition $p_1 = 0.9$ and different noise levels p_2.

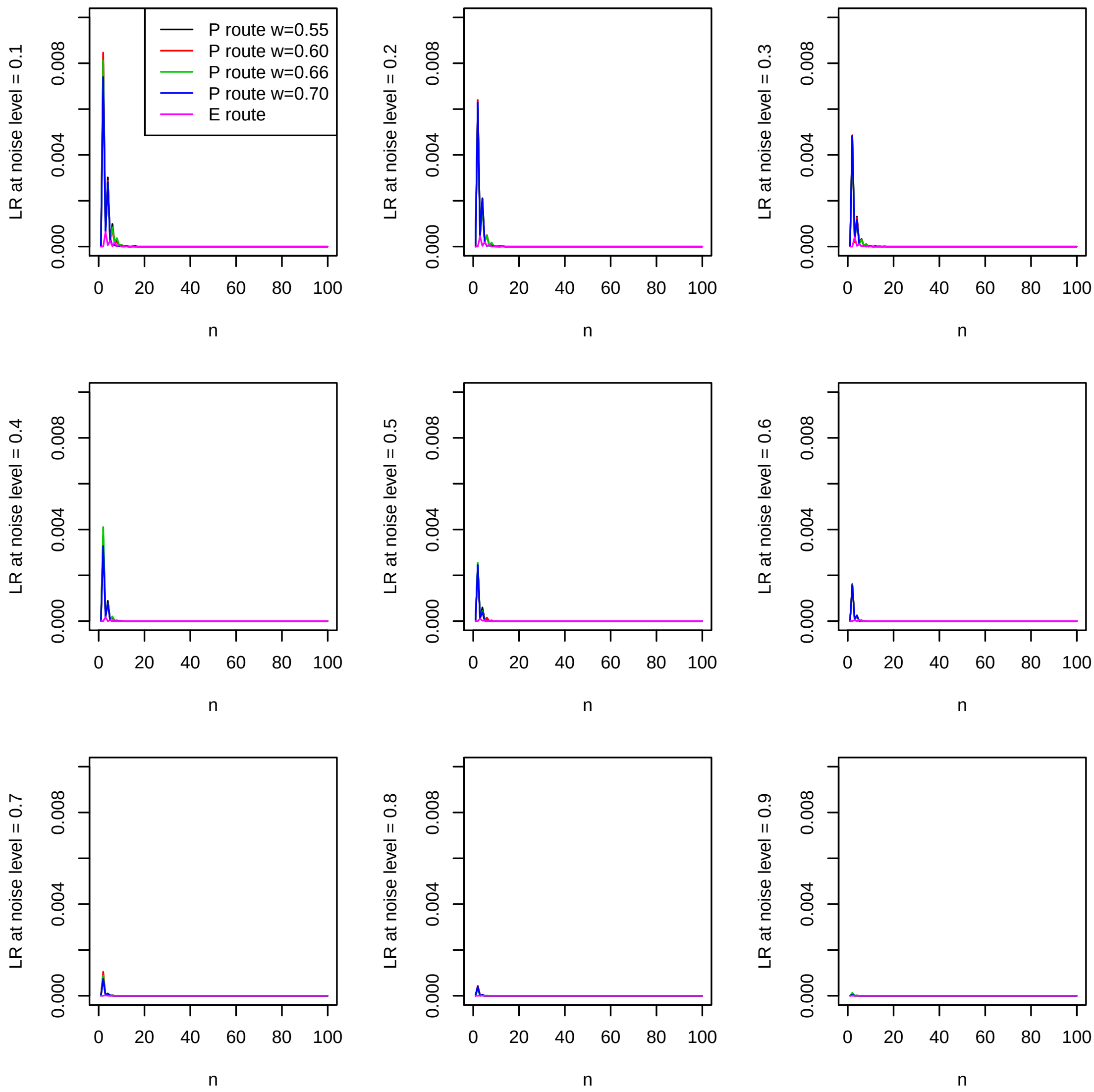

Figure 4.18: LR of routes 1.1 and 1.2 at $T = 100$ under market condition $p_1 = 0.9$ and different noise levels p_2.

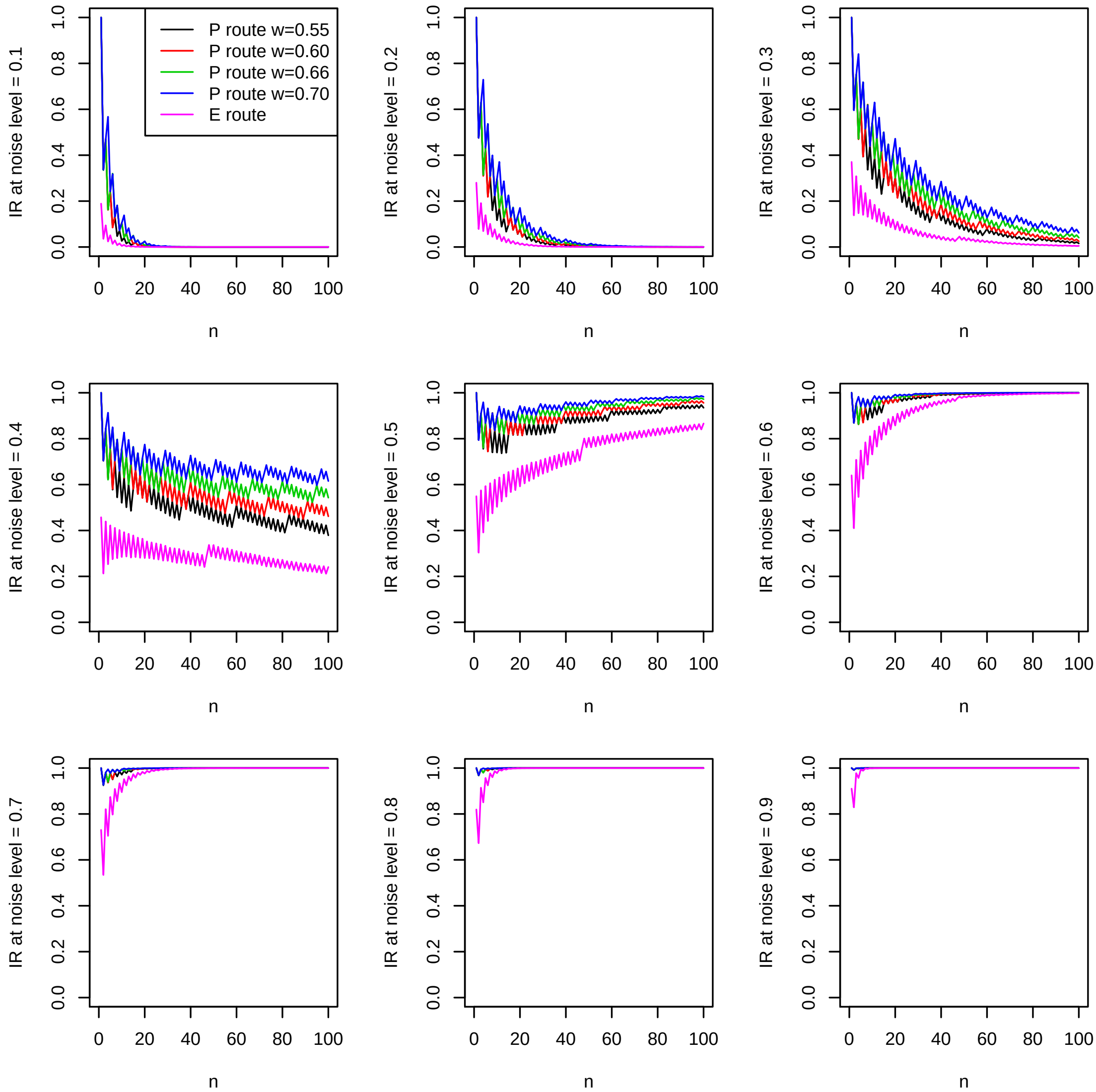

Figure 4.19: IR of routes 1.1 and 1.2 at $T = 100$ under market condition $p_1 = 0.9$ and different noise levels p_2.

From Figure 4.16, under market condition $p_1 = 0.9$, it could be found that route 1.2 is dominating in term of $\overline{f_i^A}$ regardless what noise level is presented in the data which is as expected as route 1.2 emphasize on achieve maximum present value payoff in the long run. Although from Figure 4.18, one may argue route 1.1 has worse performance in loss rate when the noise level is low, one should notice that the under market condition $p_1 = 0.9$, the magnitude of the loss ratio difference between route 1.1 and route 1.2 is less than 0.008 which is extremely small. And it is observed from other simulation results for market $p_1 \in (0.1, 0.5)$ under different noise levels, route 1.1 is better at risk control in loss rate R_{lr}^A than route 1.2. Therefore, in essence, the risk control effort of route 1.1 is significant when market has no trend or in favor of short selling the asset and less noticeable when the market is in favor of buying the asset. Overall, from Figures 4.16-4.19, one could notice that, as noise level increase, both trading routes' inaction rate R_{ir}^A increase, resulting in similar trading outcomes in the high level noise region.

4.3 Overall review of NPI asset trading simulation

From Section 4.2, one could reach the following conclusion:

The proposed NPI trading route 1.1 and route 1.2 have decent performance under all market condition and different noise levels.

Both trading routes are able extract correct underlying information from the data effectively and efficiently and take corresponding correct action different market under different market conditions. The data learning process also has moderate noise resistance when low noise level is presented. When the data is affected by high level of noise, both trading routes are able to readily recognize and stop taking any non sensible action.

Under no noise or low noise condition, given sufficient data, throughout all different market conditions, route 1.1 has better risk control on loss rate while route 1.2 is able to achieve higher average present value payoff.

Chapter 5

Application of NPI method in European option trading

In this chapter, under binomial tree model, considering the financial object European call option $\Lambda_c(A_T, K)$ and European put option $\Lambda_p(A_T, K)$ defined in Chapter 2, we apply NPI to learn the information from historical data and induced imprecise probability space on the underlying asset price $A_T(S_T)$. Based on the induce NPI imprecise probability space and using CRR non-arbitrage price as current market, two NPI European call option trading routes and two NPI European put option trading routes are proposed. Simulations are subsequently conducted to evaluate the trading routes' performance. One should notice the crucial point in this chapter is that we admit the non arbitrage price derived by the CRR model. The non arbitrage price is used as the current market price in the simulations. Also, the formulation of all trading routes in this chapter involves using the non arbitrage price as current market price and NPI imprecise probability or expectation. This is an important difference from He et al.'s work [27,28] where they use NPI expectation as an alternative option pricing model and investigate the trading result between CCR believer and NPI believer under different market conditions.

5.1 NPI method in European call option trading

In this section, we apply NPI method in European call option trading. A call option trading scenario is firstly specified. Under this scenario, two NPI call option trading routes are proposed. Subsequently, simulations are conducted to evaluate the performance of the proposed European call option trading routes by five performance indexes.

5.1.1 Call option trading Scenario setting

Consider the scenario: one is allowed to long or short the one unit of call option with strike price K and maturity date T at price $\Lambda_c^{\mathbb{Q}}(a_0, K)$ in time 0. Also, one is allowed to invest or borrow $\Lambda_c^{\mathbb{Q}}(a_0, K)$ with risk free interest rate r. Whatever position one enters, one has to keep the position for time length T and is obligated to close all risk position at time T (One is allowed to buy, sell or short sell the asset at price A_T for closing the risk position in time T). How should one, who is a NPI imprecise probability believer, without using any of his or her capital, make one's decision in trading to maximize one's capital gain in present value probabilistically or expectationally at time T? (Assume one's capital is able to cover any potential loss)?

In the scenario, the key points to emphasize are: fixed entering position time point, fixed closing position time point, one single call is available to long or short.

One is interested in the call option payoff $\Lambda_c(A_T, K) = (A_T(S_T) - K)^+$ at time T. Since $\Lambda_c(A_T, K) = (A_T - K)^+$ is monotonically increasing function of A_T and $A_T(\cdot)$ is monotonically increasing function S_T. Thus $\Lambda_c(A_T(S_T), K)$ is monotonically increasing function of S_T. One therefore can compute $\underline{E}(\Lambda_c(A_T, K))$ and $\overline{E}(\Lambda_c(A_T, K))$ by construct $\underline{p}_{\Lambda_c(A_T,K)}(\cdot)$,and $\overline{p}_{\Lambda_c(A_T,K)}(\cdot)$ using Formulas 2.2.29 and 2.2.30.

A NPI believer, who prefer to use imprecise probability operator $\underline{p}$ and $\overline{p}$ could use following European call option trading route in this scenario at time 0:

Set threshold value $0.5 < w < 1$. From NPI setting, one has $0 < \underline{p}((A_T(S_T) - K)^+ > B(T)\Lambda_c^{\mathbb{Q}}(a_0, K)) < \overline{p}((A_T(S_T) - K)^+ > B(T)\Lambda_c^{\mathbb{Q}}(a_0, K)) < 1$ if $S_T \subsetneq \mathbb{N}_0^T$ and

$S_T \neq \emptyset$.

Imprecise probability European call trading route 2.1:

$$\begin{cases} \text{Borrow cash } \Lambda_c^{\mathbb{Q}}(a_0, K) \text{ and buy the call option at time 0,} \\ \text{Exercise the option to buy the asset with price K at time } T \text{ and sell the asset} \\ \text{at market price } A_T \text{ if } A_T > K \text{, return cash } \Lambda_c^{\mathbb{Q}}(a_0, K)B(T) \text{ to the lender at time } T \\ \text{if: } \underline{p}[(A_T(S_T) - K)^+ > B(T)\Lambda_c^{\mathbb{Q}}(a_0, K)] > w \\ \\ \text{Short the call option at time 0 for } \Lambda_c^{\mathbb{Q}}(a_0, K) \text{, invest } \Lambda_c^{\mathbb{Q}}(a_0, K) \text{ for risk free rate } r \\ \text{close all the position at time } T \text{ if } \underline{p}[(A_T(S_T) - K)^+ < B(T)\Lambda_c^{\mathbb{Q}}(a_0, K)] > w \\ \\ \text{No action if none of above satisfied} \end{cases}$$

Motivation behind NPI imprecise probability European call option trading route 2.1

Consider the event $(A_T(S_T) - K)^+ > B(T)\Lambda_c^{\mathbb{Q}}(a_0, K)$ that the payoff of the call option at time T is greater than the interest $B(T)\Lambda_c^{\mathbb{Q}}(a_0, K)$ generated at time T by borrowing $\Lambda_c^{\mathbb{Q}}(a_0, K)$ at risk free rate r at time 0. If the lower probability of this event is greater than the threshold value w ($w > 0.5$), one would prefer to buy this call option and expect to earn more than $B(T)\Lambda_c^{\mathbb{Q}}(a_0, K)$ from the call option payoff in future time T.

On the contrary, consider the event $(A_T(S_T) - K)^+ < B(T)\Lambda_c^{\mathbb{Q}}(a_0, K)$ that the payoff the call option at future time T is less than interest generated by investing $\Lambda_c^{\mathbb{Q}}(a_0, K)$ at time 0. If the lower probability of this event is greater than threshold value w ($w > 0.5$), then one would expect the payoff of call option more likely to be less than the interest generated by investing $\Lambda_c^{\mathbb{Q}}(a_0, K)$ at time 0. Thus one would prefer to short sell the call option and invest the amount $\Lambda_c^{\mathbb{Q}}(a_0, K)$ with risk free rate r.

If none of above conditions are satisfied, one would better off take no action as the lower probability of the desirable event is not high enough for one to make a confident decision.

One can show that only one of actions could be taken in Route 2.1: Using inequality and conjugacy property of imprecise probability, one could know:

if $\underline{p}[(A_T(S_T)-K)^+ > B(T)\Lambda_c^{\mathbb{Q}}(a_0,K)] > w$

then, by congugacy property

$1-\overline{p}[(A_T(S_T)-K)^+ < B(T)\Lambda_c^{\mathbb{Q}}(a_0,K)] > w$

by imprecise probability inequality

$\underline{p}[(A_T(S_T)-K)^+ < B(T)\Lambda_c^{\mathbb{Q}}(a_0,K)] < \overline{p}[(A_T(S_T)-K)^+ < B(T)\Lambda_c^{\mathbb{Q}}(a_0,K)] < 1-w < w$

Thus, only one action could be taken in the imprecise probability European call option trading route.

Under the presetting scenario, a NPI believer, who prefer to use imprecise expectation operator $\underline{E}$ and $\overline{E}$ could use following European call trading route at time 0:

Imprecise expectation European call trading route 2.2:

$$\begin{cases} \text{Borrow cash } \Lambda_c^{\mathbb{Q}}(a_0,K) \text{ and buy the call option at time 0,} \\ \text{Exercise the option to buy the asset with price K at time } T \text{ and sell the asset} \\ \text{at market price } A_T \text{ if } A_T > K \text{, return cash } \Lambda_c^{\mathbb{Q}}(a_0,K)B(T) \text{ to the lender at time } T \\ \text{if: } \underline{E}[(A_T-K)^+] > B(T)\Lambda_c^{\mathbb{Q}}(a_0,K) \\ \\ \text{Short the call option at time 0 for } \Lambda_c^{\mathbb{Q}}(a_0,K) \text{, invest } \Lambda_c^{\mathbb{Q}}(a_0,K) \text{ for risk free rate } r \\ \text{close all the position at time } T \text{ if } \overline{E}[(A_T-K)^+] < B(T)\Lambda_c^{\mathbb{Q}}(a_0,K) \\ \\ \text{No action if none of above satisfied} \end{cases}$$

Motivation behind NPI imprecise expectation call option trading route 2.2
When the lower expectation of call option payoff $\underline{E}[(A_T-K)^+]$ at future time T is greater than the return $B(T)\Lambda_c^{\mathbb{Q}}(a_0,K)$ generated at time T by borrowing $\Lambda_c^{\mathbb{Q}}(a_0,K)$ at time 0 with risk free interest rate r, one would prefer to borrowing $\Lambda_c^{\mathbb{Q}}(a_0,K)$

at time 0 and buy the call option and expect to receive at least the amount of $\underline{E}[(A_T - K)^+] - B(T)\Lambda_c^{\mathbb{Q}}(a_0, K)$ at time T.

If the upper expectation of call option payoff $\overline{E}[(A_T - K)^+]$ at future time T is less than the current call option non-arbitrage price $\Lambda_c^{\mathbb{Q}}(a_0, K)$, one would rationally choose to short the call option and invest the money received into risk free rate r at time 0, expecting receive at least $B(T)\Lambda_c^{\mathbb{Q}}(a_0, K) - \overline{E}[(A_T - K)^+]$ at time T by close all the position.

It is also easy to show that only one of actions could be taken in Route 2.2: In the imprecise probability framework, one has $\underline{E}[(A_T - K)^+] \leq \overline{E}[(A_T - K)^+]$ and therefore $\underline{E}[(A_T - K)^+] > B(T)\Lambda_c^{\mathbb{Q}}(a_0, K)$ and $\overline{E}[(A_T - K)^+] < B(T)\Lambda_c^{\mathbb{Q}}(a_0, K)$ could not be satisfied at the same time.

5.1.2 Simulation of call option trading in NPI Bernoulli model

In this section, we use simulation to study the performance of two proposed NPI European call option trading routes in the prescribed scenario setting.

We only present simulation results with following valued predefined parameters r,u,d and a_0, K. Other values of predefined parameters value are also simulated; they all have similar patterns.

[Predefined parameters value for r, u, d, a_0 and K] We use the same predefined parameter value for r, u, d, a_0 as in the asset trading chapter and call option strike price K is set at $K = 103$

All the trading routes are simulated 100,000 times using the statistical software R version 3.5.1. The data generating process of the underlying asset in this section is the same as the previous chapter and thus will not be repeatedly stated here.

Performance evaluation function f_i^C

The performances of NPI European call trading routes are measured by five statistics of the present value pay-off function $f_i^C(n, T, i)$ in 100000 simulations. $f_i^C(n, T, i)$ is defined as follow:

$$f_i^C(n,T,i) = \begin{cases} (A_T(S_T)-K)^+B(T)^{-1} - \Lambda_c^{\mathbb{Q}}(a_0,K) & \text{if first action of the trading route is taken} \\ \Lambda_c^{\mathbb{Q}}(a_0,K) - (A_T(S_T)-K)^+B(T)^{-1} & \text{if second action of the trading route is taken} \\ 0 \quad \text{if no action} \end{cases}$$

where the inputs:

n is the length of historical asset price data one could learn;

T is the future time that the this function is evaluate;

$i \in (1, 100000)$ is the index of that particular simulation trial.

Five performance statistics of this function measure from 100000 simulations are:

Average present value payoff $\overline{f_i^C} = \dfrac{\sum_i f_i^C}{100000}$ Win-loss ratio $R_{wl}^C = \dfrac{|\{i: f_i^C > 0\}|}{|\{i: f_i^C < 0\}|}$

Win rate $R_{wr}^C = \dfrac{|\{i: f_i^C > 0\}|}{100,000}$ Loss rate $R_{lr}^C = \dfrac{|\{i: f_i^C) < 0\}|}{100,000}$

Inaction rate $R_{ir}^C = \dfrac{|\{i: f_i^C = 0\}|}{100,000}$

Sample simulation trials of different call option trading routes given precise or imprecise data

Several simulation trials are provided to illustrate how each call option trading route work in the simulation process.

Simulation trial 1 Underlying market condition $p = 0.2$ (For the investor, this information is hidden), one observes following precise data of the underlying asset price in past 7 time stages

Time stage	-7	-6	-5	-4	-3	-2	-1
Data ($p = 0.2$)	0	0	0	0	1	1	0
Equivalently $(n, j) = (7, 2)$							

With predefined parameters value, one needs to decide whether or not enter a risk position of a call option of which the mature time is at time $T = 7$. By CRR pricing model, $q = \frac{e^r - d}{u - d} = 0.5044$ and current market price $\Lambda_c^{\mathbb{Q}}(a_0, K)$ is

$$\begin{aligned}
\Lambda_c^{\mathbb{Q}}(a_0, K) &= B(T)^{-1} \sum_{S_T=0}^{T} \binom{T}{S_T} (A_0 u^{S_T} d^{T-S_T} - K)^+ q^{S_T} (1-q)^{T-S_T} \\
&= B(7)^{-1} \sum_{S_T=0}^{7} \binom{7}{S_T} (100 * 1.03^{S_T} (\frac{1}{1.03})^{7-S_T} - 103)^+ q^{S_T} (1-q)^{7-S_T} \\
&= 2.0094
\end{aligned}$$

If one uses route 2.1 (imprecise probability trading route) and set threshold value $w = 0.7$. One firstly finds out m such that $(A_7(m) - K)^+ = B(T)\Lambda_c^{\mathbb{Q}}(a_0, K)$. $m \approx 4.319$, One then find out $m_1 = \lceil m \rceil = 5$, $m_2 = \lfloor m \rfloor = 4$ and calculate $\underline{p}_{(7,2)}(S_7 \geq m_1) = 0.0512 < w$ and $\underline{p}_{(7,2)}(S_7 \leq m_2) = 0.8569 > w$, thus one will take second action of route 2.1 in this case, namely, one will short the call option at time 0 for 2.0094, invest 2.0094 for risk free rate 0.003 and close all the position at time 7

If one uses route 2.2 (imprecise expectation trading route), one will find out $\underline{E}_{(7,2)}[(A_T - K)^+] = 0.4350 < \Lambda_c^{\mathbb{Q}}(a_0, K)B(7) = 1.9618$ and $\overline{E}_{(7,2)}(A_7) = 1.3127 < \Lambda_c^{\mathbb{Q}}(a_0, K)B(7) = 1.9618$. Thus, one will take second action of route 2.2 and execute the same strategy as one uses route 2.1.

Simulation trial 2 Underlying market condition $p = 0.5$, one observes following data of the underlying asset price in past 7 time stages

Time stage	-7	-6	-5	-4	-3	-2	-1
Data ($p = 0.5$)	0	1	0	0	1	0	1
Equivalently $(n, j) = (7, 3)$							

One needs to decide whether or not enter a risk position of a call option which expired at future 7 time units. If one uses route 2.1 (imprecise probability trading route) and set threshold value $w = 0.6$. With the predefined parameters value, current market price $\Lambda_c^{\mathbb{Q}}(a_0, K)$ is still 2.0094. The value m such that $(A_7(m) - K)^+ = B(T)\Lambda_c^{\mathbb{Q}}(a_0, K)$ is still $m \approx 4.319$. One then still find out $m_1 = \lceil m \rceil = 5$, $m_2 = \lfloor m \rfloor = 4$ calculate $\underline{p}_{(7,3)}(S_7 \geq m_1) = 0.1430 < w$ and $\underline{p}_{(7,3)}(S_7 \leq m_2) = 0.7040 > w$, therefore one uses route 2.1 will take the second action.

If one uses route 2.2 (imprecise expectation trading route), one will find out $\underline{E}_{(7,3)}[(A_T - K)^+] = 1.312 < \Lambda_c^{\mathbb{Q}}(a_0, K)B(7) = 1.9618$ and $\overline{E}_{(7,3)}(A_7) = 2.987 > \Lambda_c^{\mathbb{Q}}(a_0, K)B(7) = 1.9618$. Thus, one would take no action if one uses route 2.2 in this case.

Simulation trial 3 Underlying market condition $p_1 = 0.2$, and noise level $p_2 = 0.2$, one observes following data of the underlying asset price in past 7 time stages.

Time stage	-7	-6	-5	-4	-3	-2	-1
Data ($p_1 = 0.2$, $p_2 = 0.2$)	0	0	{0,1}	0	{0,1}	0	0
Equivalently $[n, J] = [7, (0, 2)]$							

One again needs to make decision whether or not enter a risk position of the option. If one uses route 2.1 (imprecise probability trading route) and set threshold value $w = 0.65$. The value m such that $(A_7(m)-K)^+ = B(T)\Lambda_c^{\mathbb{Q}}(a_0, K)$ is still $m \approx 4.319$. Set $m_1 = \lceil m \rceil = 5$, $m_2 = \lfloor m \rfloor = 4$ and calculate $\underline{p}_{[7,(0,2)]}(S_7 \geq 5) = 0 < w$ and $\underline{p}_{[7,(0,2)]}(S_7 \leq 4) = 0.8569 > w$ so one will take second action of route 2.1 in this case.

If one uses route 2.2 (imprecise expectation trading route), one will find out $\underline{E}_{[7,(0,2)]}[(A_T - K)^+] = 0 < \Lambda_c^{\mathbb{Q}}(a_0, K)B(7) = 1.9618$ and $\overline{E}_{(7,(0,2))}(A_7) = 1.312 < \Lambda_c^{\mathbb{Q}}(a_0, K)B(7) = 1.9618$. Thus, one will take second action of route 2.2 in this case.

Simulation trial 4 Underlying market condition $p_1 = 0.7$, and noise level $p_2 = 0.6$, one observes following data of the underlying asset price in past 7 time stages.

Time stage	-7	-6	-5	-4	-3	-2	-1
Data ($p_1 = 0.7$, $p_2 = 0.6$)	0	{0,1}	{0,1}	1	1	{0,1}	0
Equivalently $[n, J] = [7, (2, 5)]$							

With the same situation, one uses route 2.1 (imprecise probability route) and set threshold value $w = 0.6$. One calculate $\underline{p}_{[7,(2,5)]}(S_7 \geq 5) = 0.0512 < w$ and $\underline{p}_{[7,(2,5)]}(S_7 \leq 4) = 0.2797 < w$ so one will take no action in this case.

If one uses route 2.2 (imprecise expectation trading route), one will find out $\underline{E}_{[7,(0,2)]}[(A_T - K)^+] = 0.4351 < \Lambda_c^{\mathbb{Q}}(a_0, K)B(7) = 1.9618$ and $\overline{E}_{(7,(0,2))}[(A_T - K)^+] =$

$9.4747 > \Lambda_c^{\mathbb{Q}}(a_0, K)B(7) = 1.9618$. Thus, one who use route 2.2 will take no action in this case.

Performance of NPI call option trading routes under average market condition given precise data available

Under the average market condition, given precise data available, the performances of NPI European call option trading routes 2.1 and 2.2 are assessed and discussed below.

Figure 5.1 plots the average present value payoff surface $\overline{f_i^C}$ of trading routes 2.1 and 2.2 for a call option with presetting parameter values and different maturity data $T \in (1, 100)$, given $n \in (1, 100)$ units of historical data available. One could confirm that for all size historical data $n \in (1, 100)$, both from Figure 5.1 that both routes 2.1 and 2.2 produce positive present value payoff in the long run under average market condition. The surface shares a similar pattern to the corresponding surface in asset trading chapter. Both trading routes 2.1 and 2.2 have a fast speed of learning in data. At present of $n = 15$ historical data, both trading routes have the excellent trading results for all call option expired in future 1 to 100 time units.

One may notice with a small amount of data available, the average present value payment surface has a "fan" shape in route 2.1 while route 2.2 does not have this phenomenon. The reason for this is similar to the corresponding case in asset trading chapter. Route 2.1 is an imprecise probability trading routes which aim to minimize loss rate and avoid trading in the uncertainty situation while route 2.2 is an imprecise expectation trading routes which aim to achieve higher present value payoff in the long run. This can be further confirmed in Figure 5.2.

Figure 5.2 presents the performances of $\overline{f_i^C}$,R_{wl}^C,R_{wr}^C, R_{lr}^C, and R_{ir}^C of both trading routes for call option expired in future 100 time units. It could be observed that route 2.2 generally have greater average present value payoff for different size of historical data available. However, route 2.2 has worse loss rate R_{lr}^C and win loss ratio R_{wl}^C than route 2.1. Especially in the case where a small amount of data is presented, the loss rate R_{lr}^C difference between route 2.1 and route 2.2 is significantly higher, with loss rate R_{lr}^C of route 2.2 being around 0.3. The reason for this is that route 2.1 tends to avoid making trading when the number of data is insufficient to extract enough information about the underlying distribution. (indicated by high inaction rate R^C_{ir} in Figure 5.2 when small of data is available). In addition, with adjustment of threshold value w, route 2.1 also have better control in loss rate R^C_{lr} and win-loss ratio R^C_{wl}.

Overall, given precise data, under the average market condition, both proposed European call option trading routes are able to yield positive average present value payoff and have good performance in R^A_{wr}, R^A_{wl} and R^A_{lr}. Route 2.2 has better performance in terms of average present value payoff, while route 2.1 have better control in loss rate and win-loss ratio.

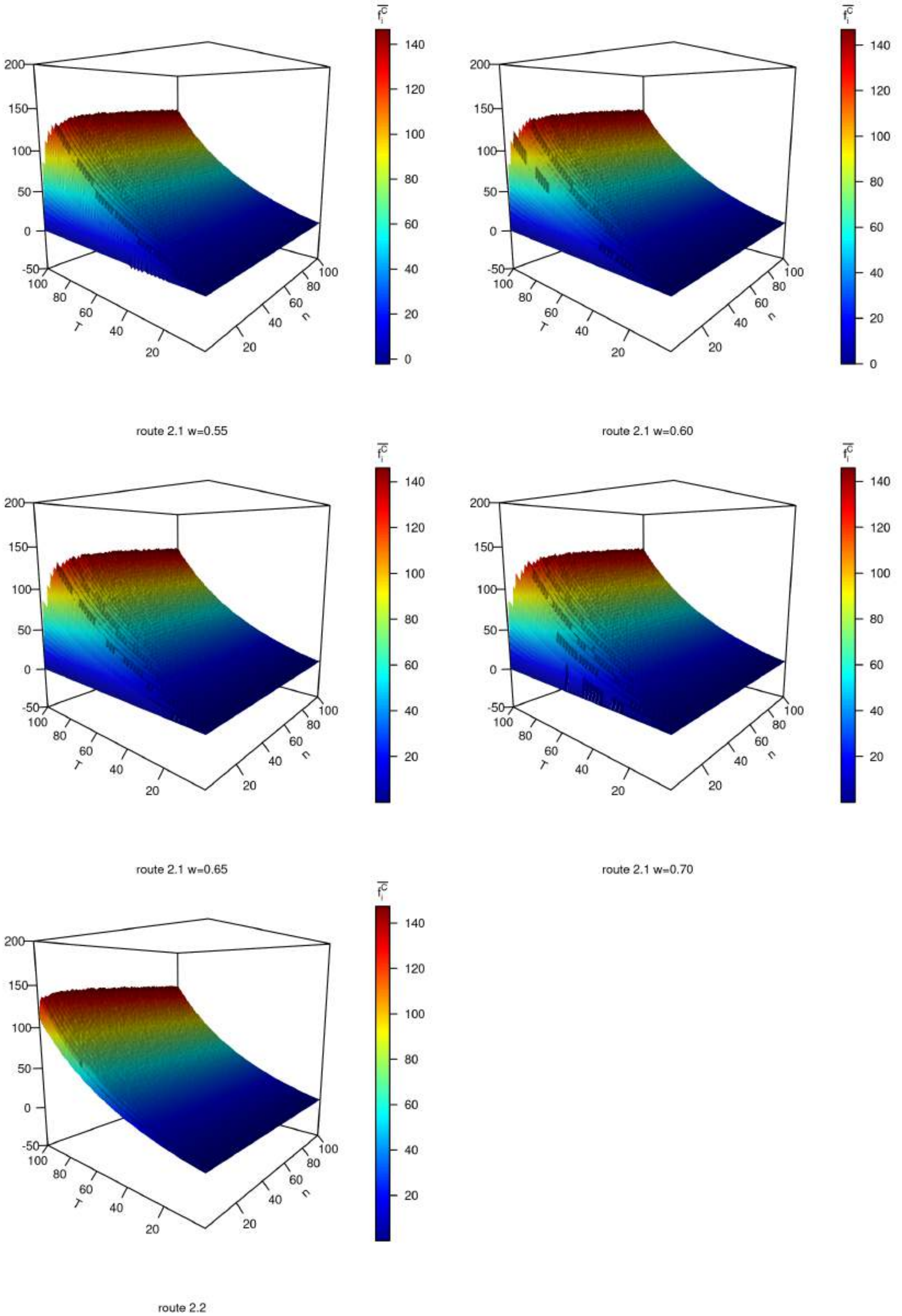

Figure 5.1: APVP of routes 2.1 and 2.2 under average market condition given precise data.

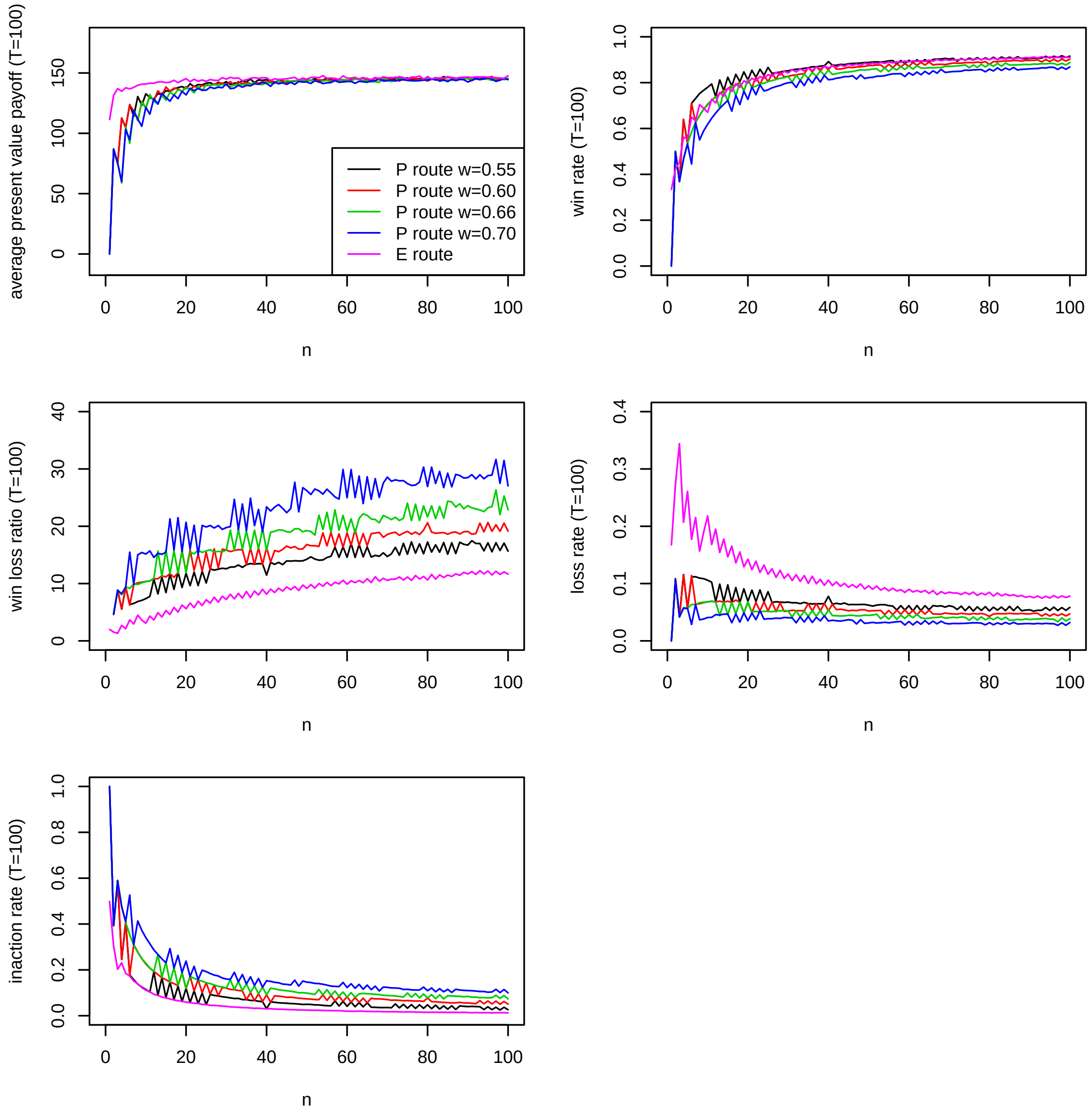

Figure 5.2: Performance comparison of routes 2.1 and 2.2 for a call option expired at time $T = 100$ under average market condition given precise data.

Performance of NPI call option trading routes under different market conditions given precise data available

Given precise data, under a specific market condition, the performance of proposed NPI European call option trading routes are further evaluated below.

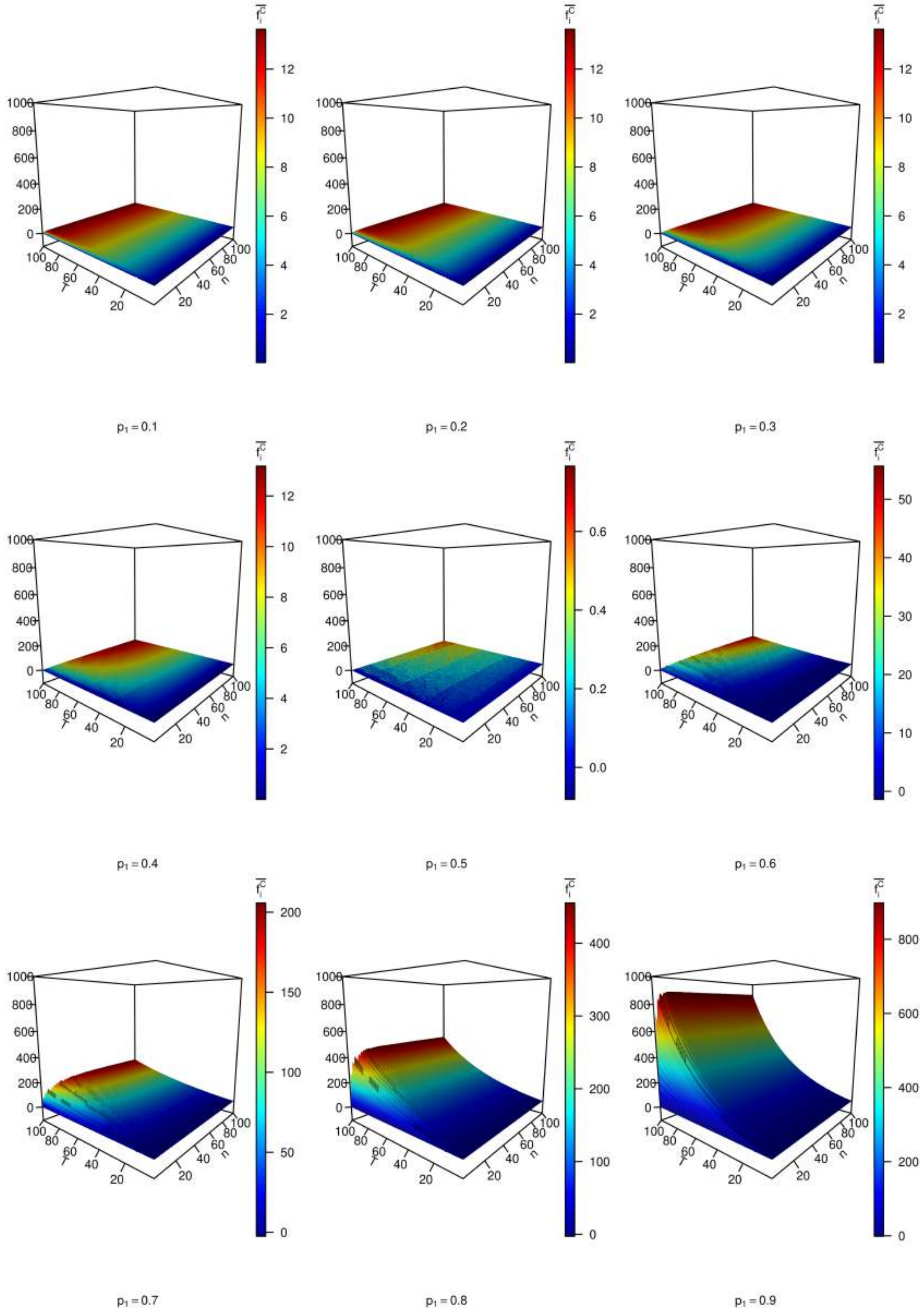

Figure 5.3: Under different market conditions, APVP of trading route 2.1 with threshold value $w = 0.6$.

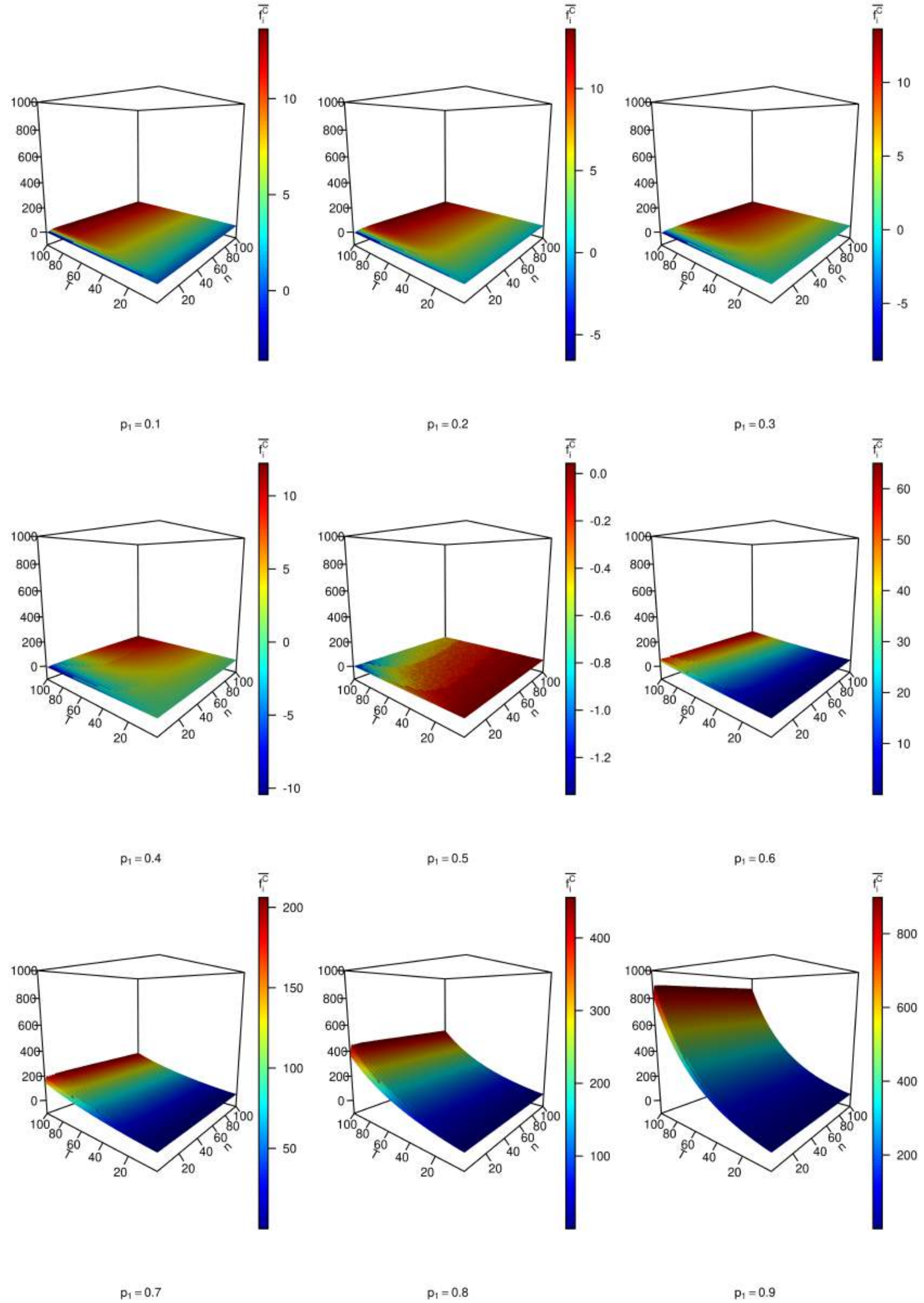

Figure 5.4: Under different market conditions, APVP of trading route 2.2.

Given precise data, Figure 5.3 and Figure 5.4 demonstrate the average present value payoff of routes 2.1 and route 2.2 under different market conditions.

It could be observed both trading routes 2.1 and 2.2 are well performed under extreme market condition $p \in (0.1, 0.3) \cup (0.7, 0.9)$, indicating both routes have the ability to recognize the underlying market condition. The simulation used the CCR European call option non-arbitrage price as market price. For an European call option which expired at time $T = 100$ with presetting parameter values, the non-arbitrage price $\Lambda_c^{\mathbb{Q}}(a_0, K)$ is 13.631. Thus the maximum present value payment of call option expired in time $T \in (1, 100)$ could be achieved by shorting sell is less than 13.631. When the market is unfavorable for the underlying asset price $p \in (0.1, 0.3)$, both trading routes most of time are able to correctly execute the second action of the trading routes, entering a short position of the call option. This results in average present value payoff ranging from 0 to 13.631 given different size of data available. In contrast, when the market is favorable for the underlying asset price $p \in (0.7, 0.9)$, both trading routes are able to correctly execute the first action, entering the long position of the call option.

It should be noticed that route 2.1 is better at avoiding making losses in trading. One could see that both trading routes have the worst performance in $\overline{f_i^C}$ when the market condition for the underlying asset price is relatively neutral $p \in (0.4, 0.6)$. In this case, making trading is more or less like a gamble. Route 2.1 is better at avoiding taking action in those situations and maintain positive $\overline{f_i^C}$ through those market conditions whereas route 2.2 are less prone to avoid trading in this situation, resulting some of negativity in $\overline{f_i^C}$ when a small amount of data is available.

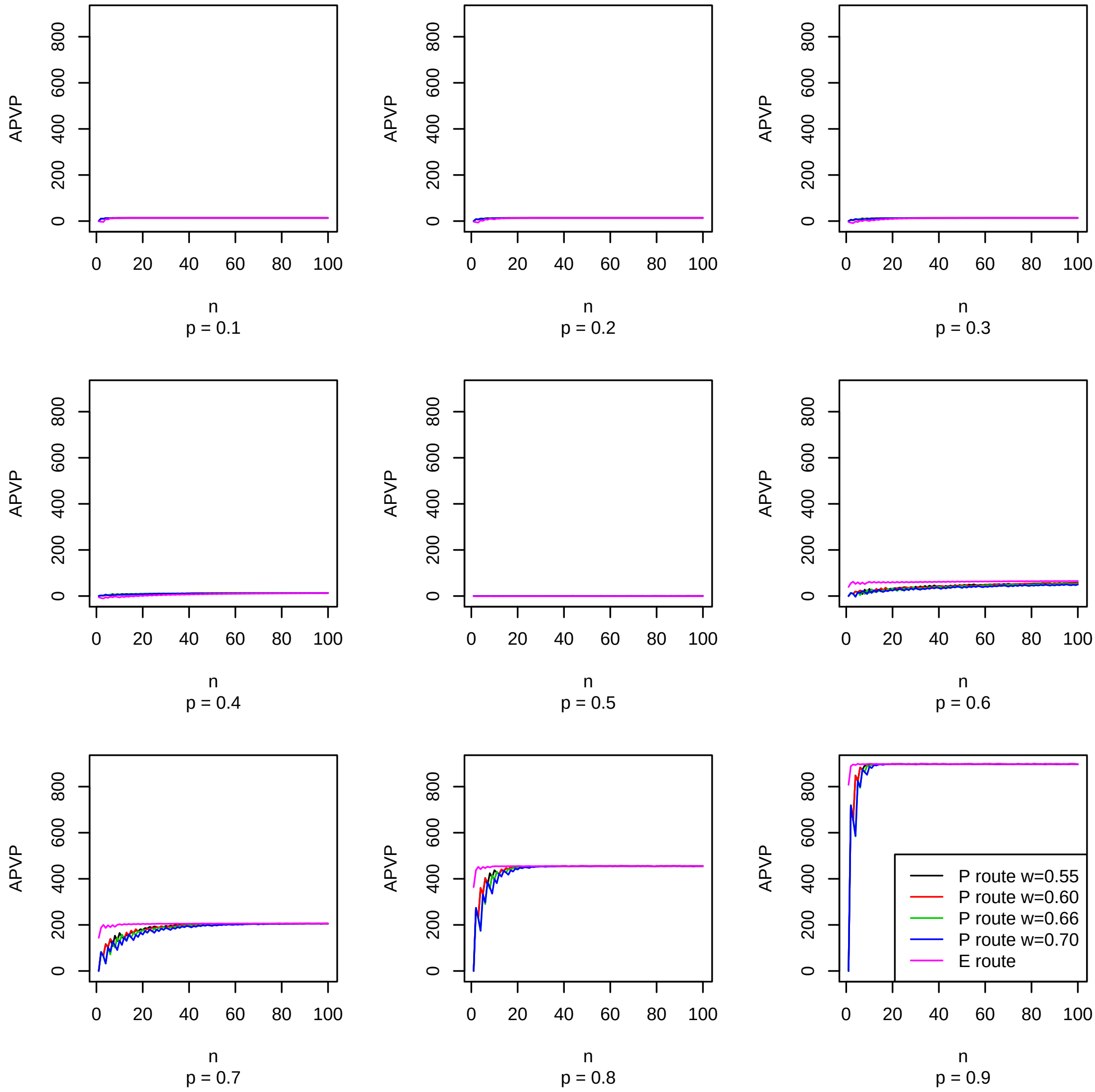

Figure 5.5: Under different market conditions, APVP of trading routes 2.1 and 2.2 for call option expired at time $T = 100$.

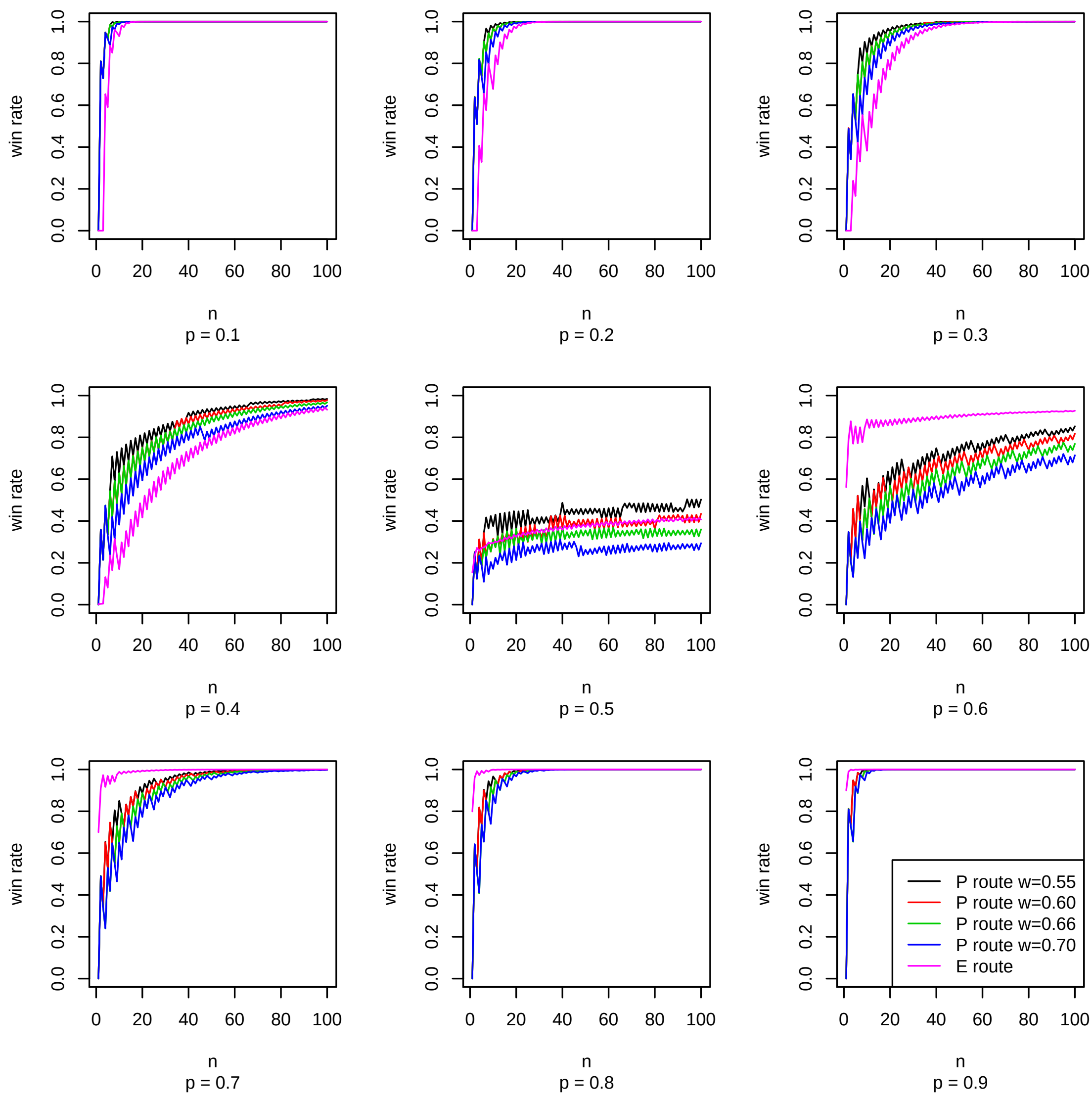

Figure 5.6: Under different market conditions, WR of trading routes 2.1 and 2.2 for call option expired at time $T = 100$.

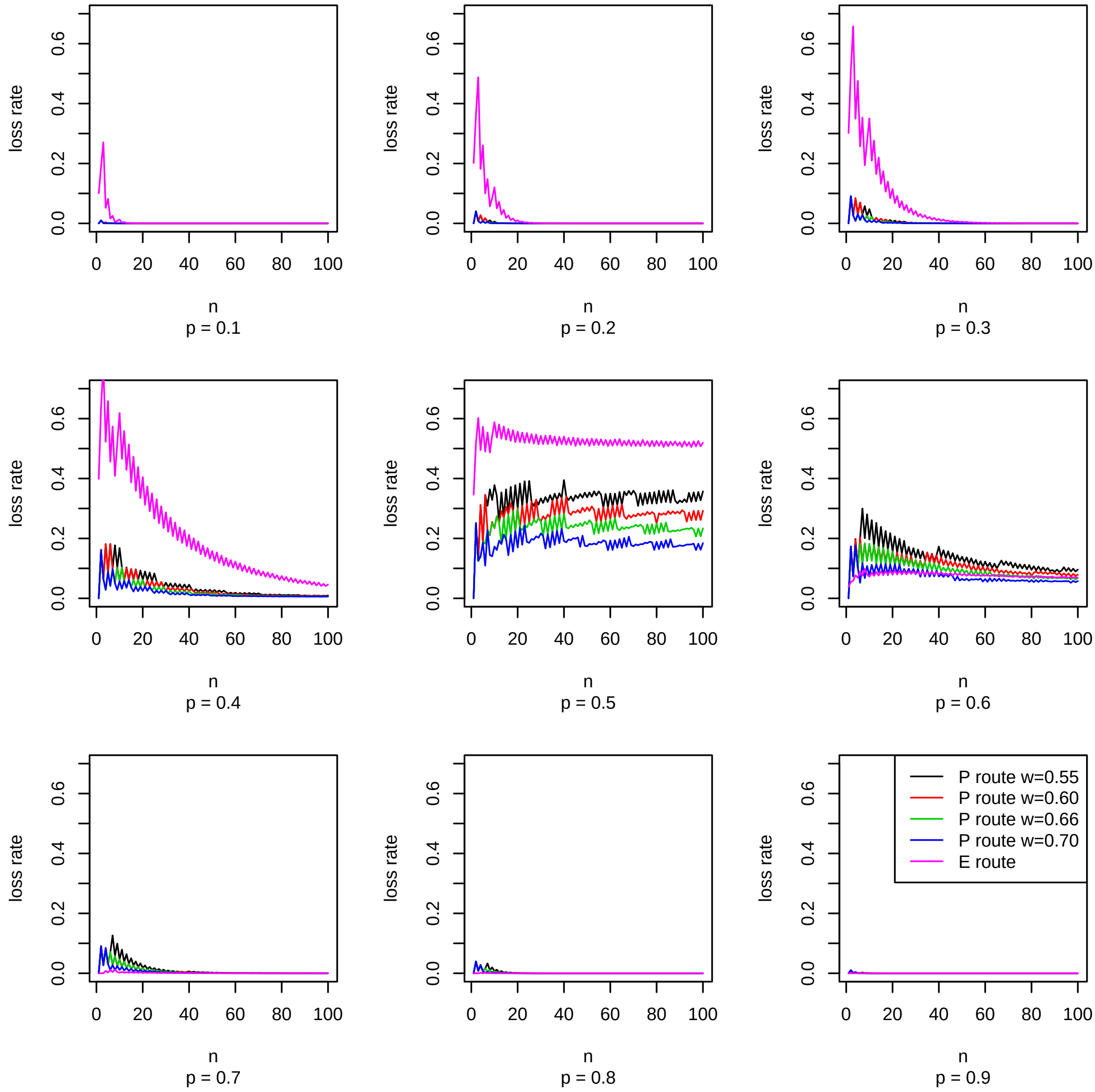

Figure 5.7: Under different market conditions, LR of trading routes 2.1 and 2.2 for call option expired at time $T = 100$.

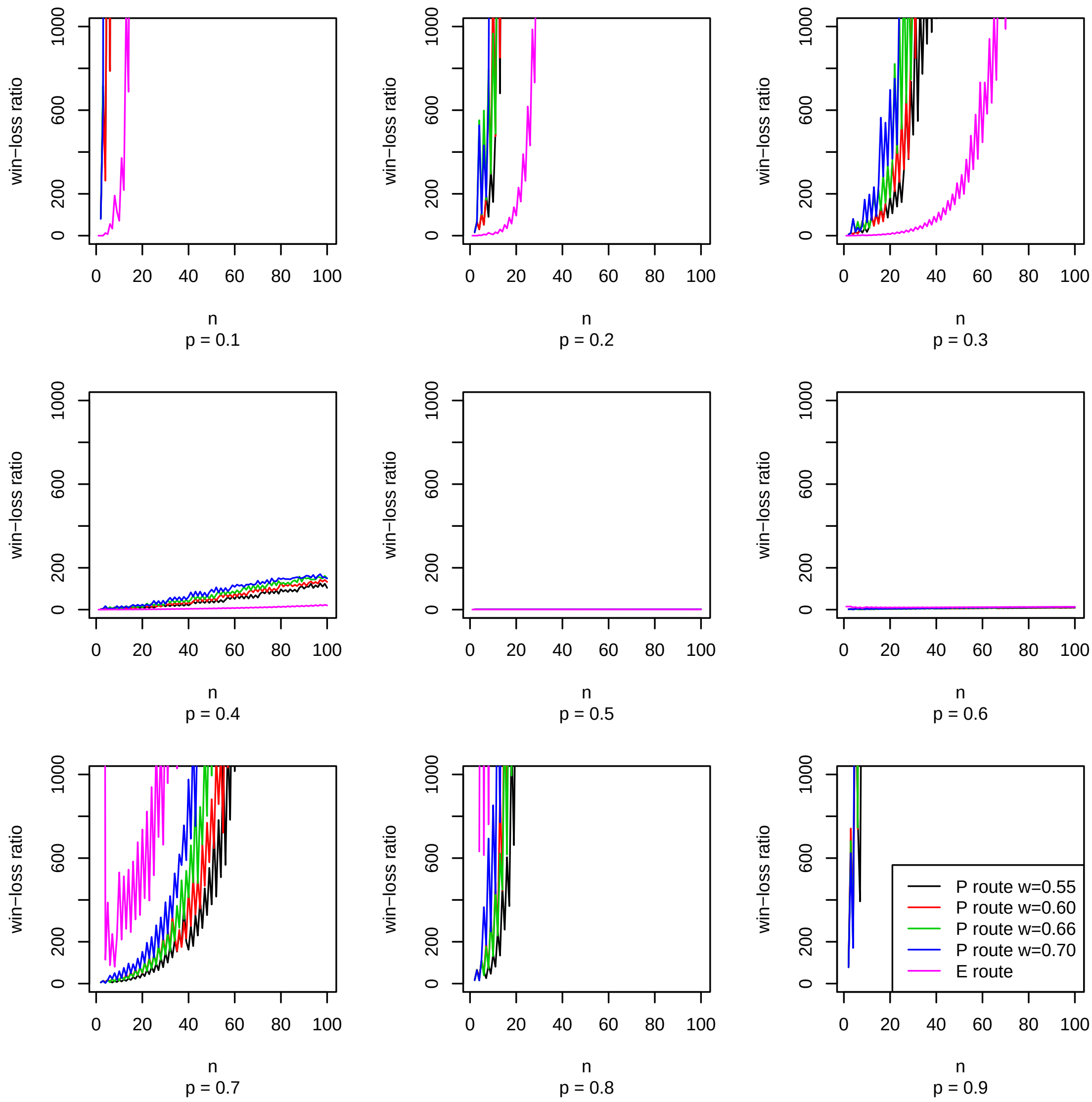

Figure 5.8: Under different market conditions, WLR of trading routes 2.1 and 2.2 for call option expired at time $T = 100$.

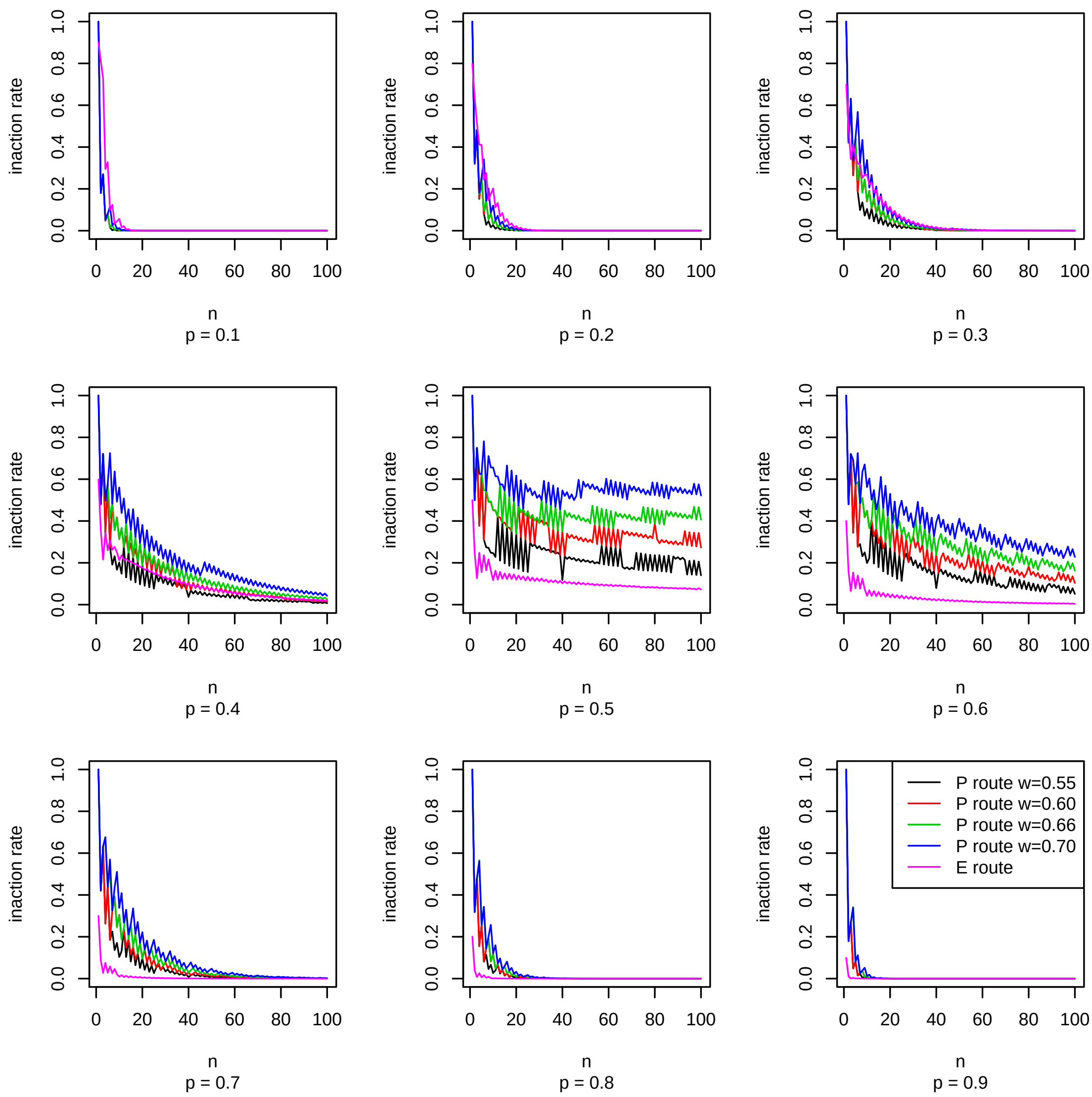

Figure 5.9: Under different market conditions, IR of trading routes 2.1 and 2.2 for call option expired at time $T = 100$.

In order to have a more detail examination of trading routes' performance, the performances index $\overline{f_i^C}$, R^C_{wr}, R^C_{lr}, R^C_{ir} for a call option expired at time $T = 100$ under different market conditions p of both trading routes are plotted in Figures 5.5-5.9.

From Figure 5.9, one could see that both trading routes indeed avoid make trading under neutral market condition $p \in (0.4, 0.6)$, indicated by higher inaction rate in the figure. Moreover, route 2.1 has strong resistance in making trading than route 2.2 in those case, which results in lower loss rate R^C_{lr} in Figure 5.7.

Overall, route 2.1 have a better win-loss ratio in most market condition markets.

Route 2.2, on the other hand, although underperformed in R^C_{lr}, R^C_{wl}, given enough data is available, it generally has higher $\overline{f^C_i}$ when market condition is two-sided. (See Figure 5.5)

Performance of NPI call option trading routes under average market condition given imprecise data available

Under the average market condition, subject to the different noise levels, the performances of NPI European call option trading routes are evaluated below.

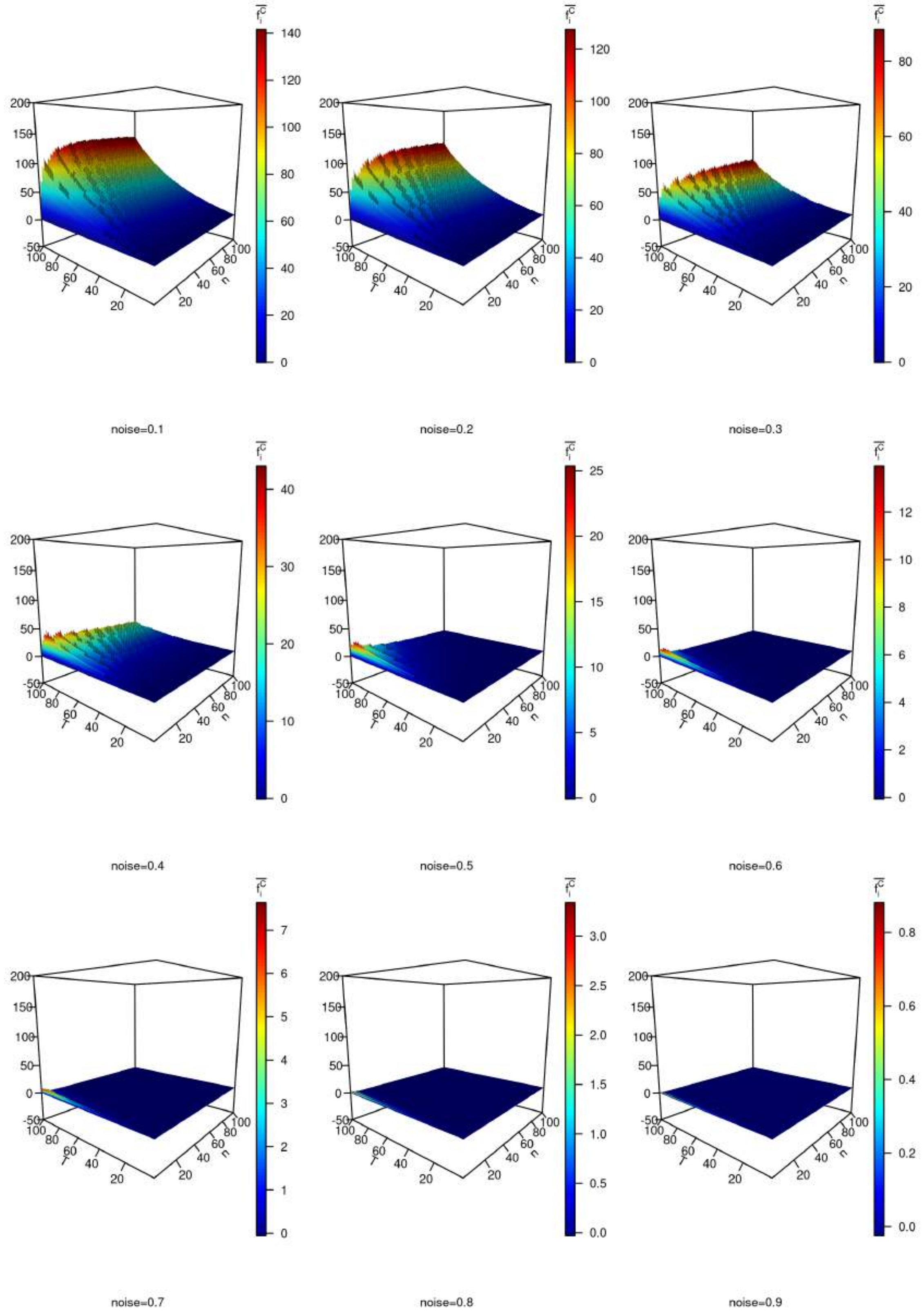

Figure 5.10: With average market condition, APVP of trading route 2.1 with threshold value $w = 0.6$ under different noise levels p_2.

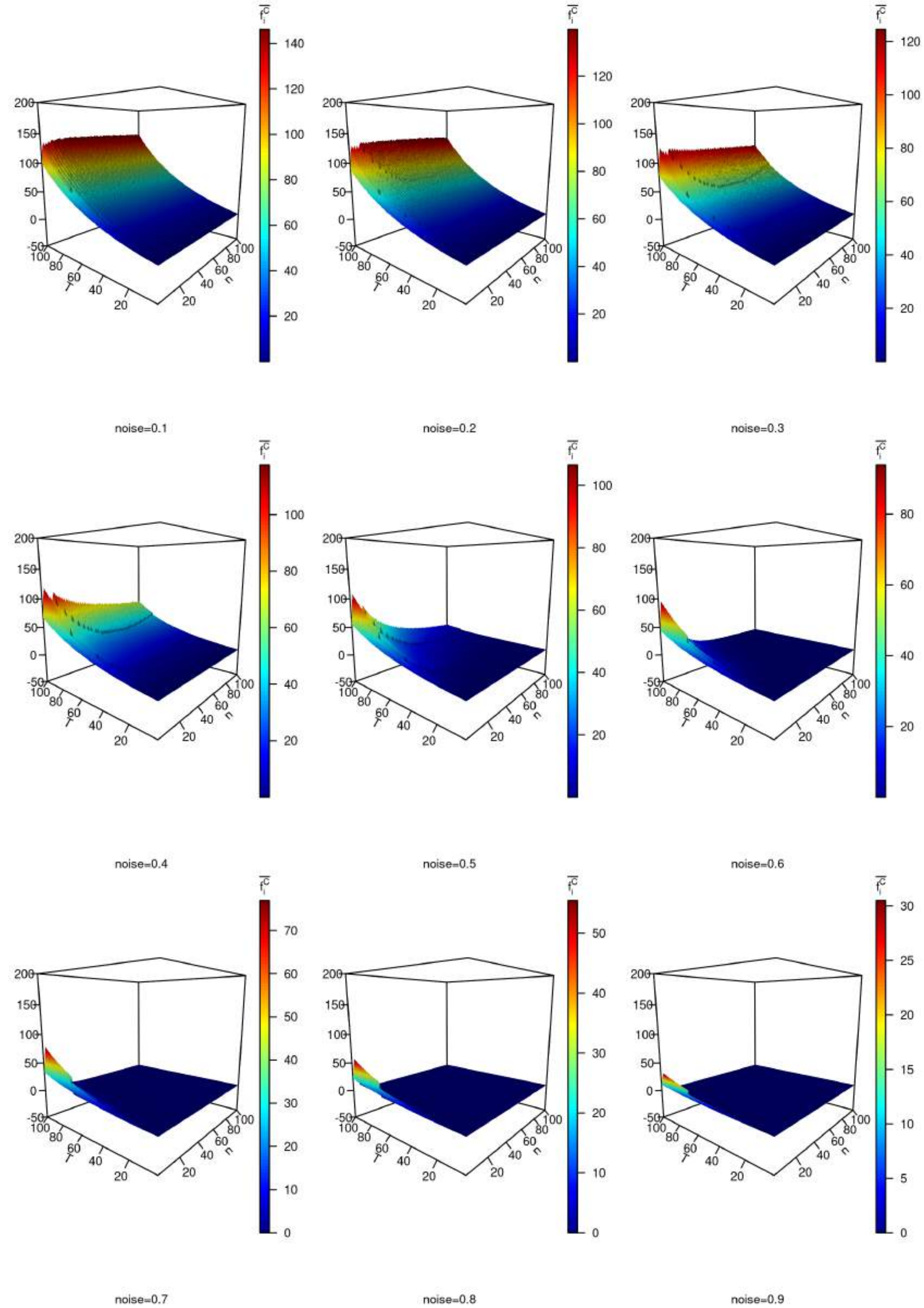

Figure 5.11: With average market condition, APVP of trading route 2.2 under different noise levels p_2.

Given n units data points, the average present value payoff $\overline{f_i^C}$ of route 2.1 with threshold value $w = 0.6$ and route 2.2 for different call options expired in 1 to 100 units of future time is plotted in Figures 5.10 and 5.11 respectively.

It could be noticed that with noise level increase in data, the $\overline{f_i^C}$ surfaces of both trading routes 2.1 and route 2.2 are decaying. This implies that both routes 2.1 and route 2.2 are able to recognize the noise level in the data. Since route 2.1 is a probability trading route which is more conservative and emphasizes more on the risk control, its speed of decaying in $\overline{f_i^C}$ is much quicker than route 2.2. This indicates that as the noise level increase, route 2.1 tends to stay more frequent in inaction than route 2.2. Nevertheless, under the average market condition, regardless of what noise level appears, both trading routes are able to yield positive average present payoff for all different combinations of the number of data points available and option expiration date. When the noise level is low, both trading routes 2.1 and route 2.2 are able to extract the correct information for the data and execute the correct action accordingly, resulting in similar shapes of surfaces $\overline{f_i^C}$ in the precise data case. (See Figure 5.1)

In order to have a more direct performance comparison of trading route 2.1 and route 2.2, $\overline{f_i^C}$, R_{wr}^C, R_{lr}^C, R_{ir}^C for a call option expired at time $T = 100$ of trading route 2.1 and route 2.2 are plotted against each other in Figures 5.12-5.15. Since as noise level increases, both trading routes 2.1 and 2.2 increase their frequency of inaction and have very low loss rate. Thus win-loss ratio R_{wr}^C is not presented.

As one has expected, the inaction rate R_{ir}^C in Figure 5.15 surged as noise level increase which confirms both trading routes 2.1 and 2.2 are able to recognize the noise level and stop making non sensible trading in the ambiguous situations. When noise level $p_2 \geq 0.5$, after sufficient enough data is gathered ($n > 62$), both trading routes stop doing any trading.

Yet when the noise level is relatively low $p_2 \in (0.1, 0.3)$, the information contained in the data is still clear. After a certain amount of data gathered, both trading routes are able to learn useful information about the underlying market condition and thus maintain moderate trading rate. It could be observed from Figure 5.14 and Figure 5.15 that route 2.1 has better performance in avoiding making losses in trading. In contrast, although route 2.2 is a more risky trading route, it, however, has higher

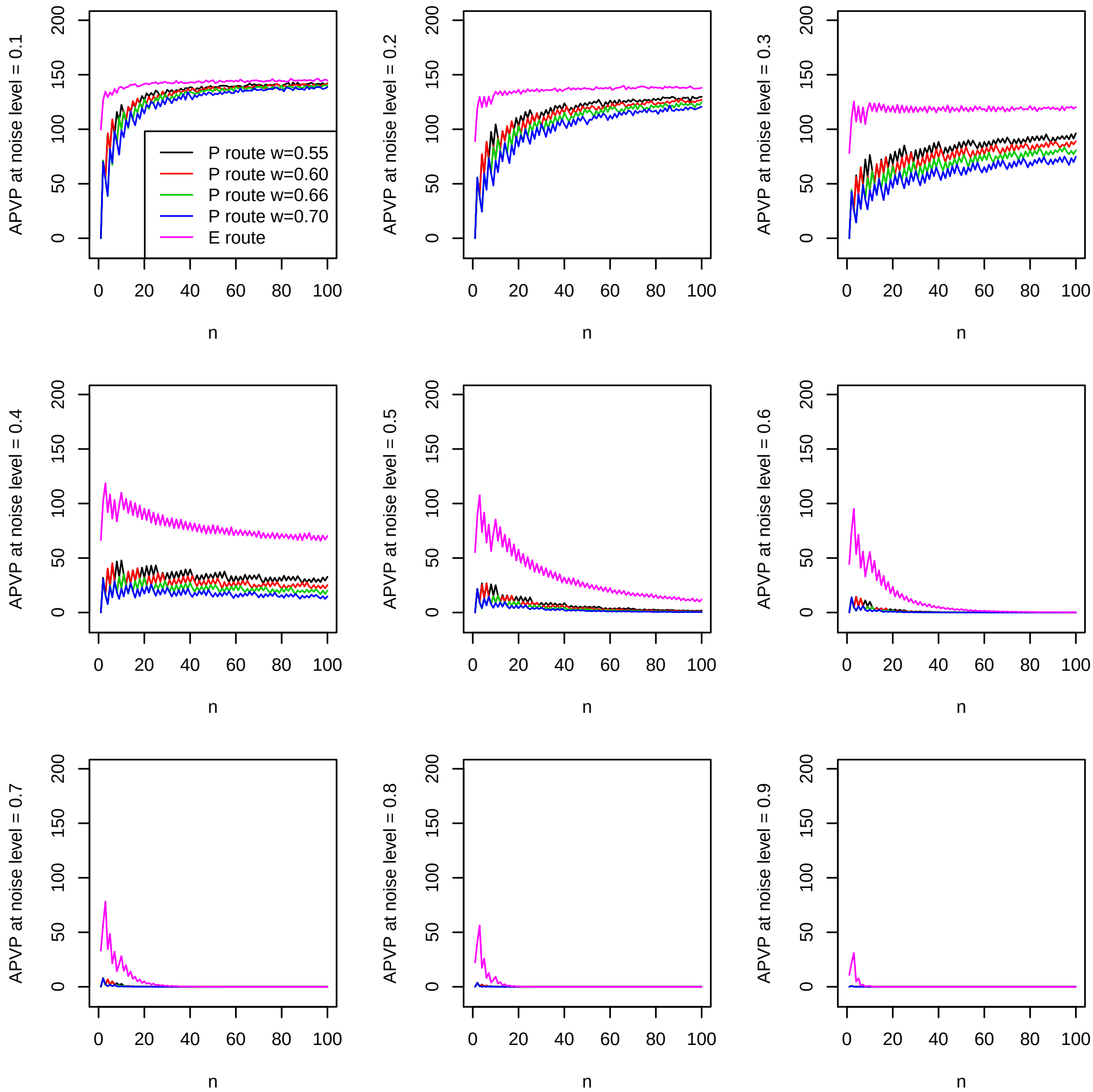

Figure 5.12: With average market condition, APVP of both trading routes 2.1 and 2.2 under different noise levels p_2 for a call option expired data at $T = 100$.

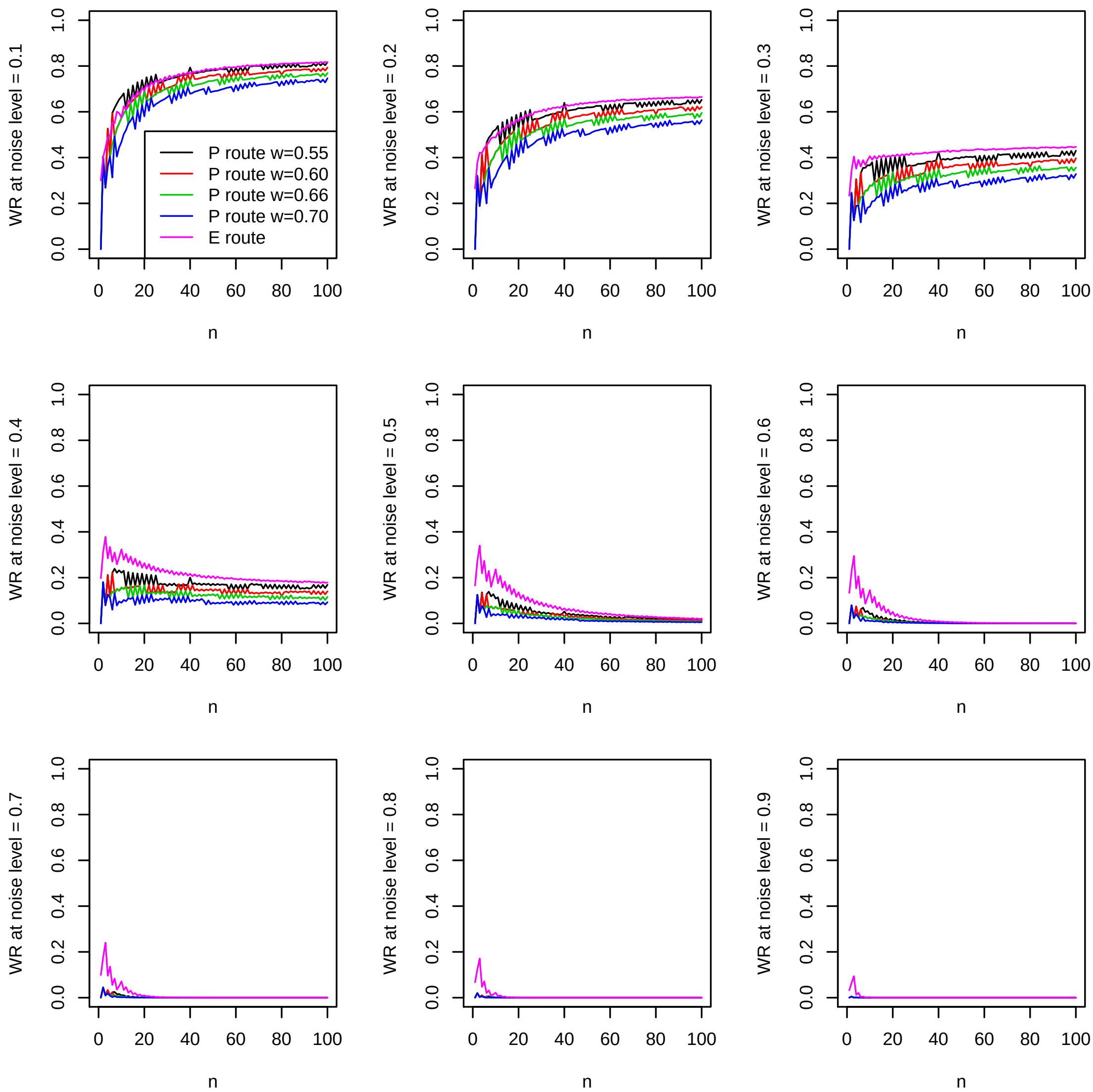

Figure 5.13: With average market condition, WR of both trading routes 2.1 and 2.2 under different noise levels p_2 for a call option expired data at $T = 100$.

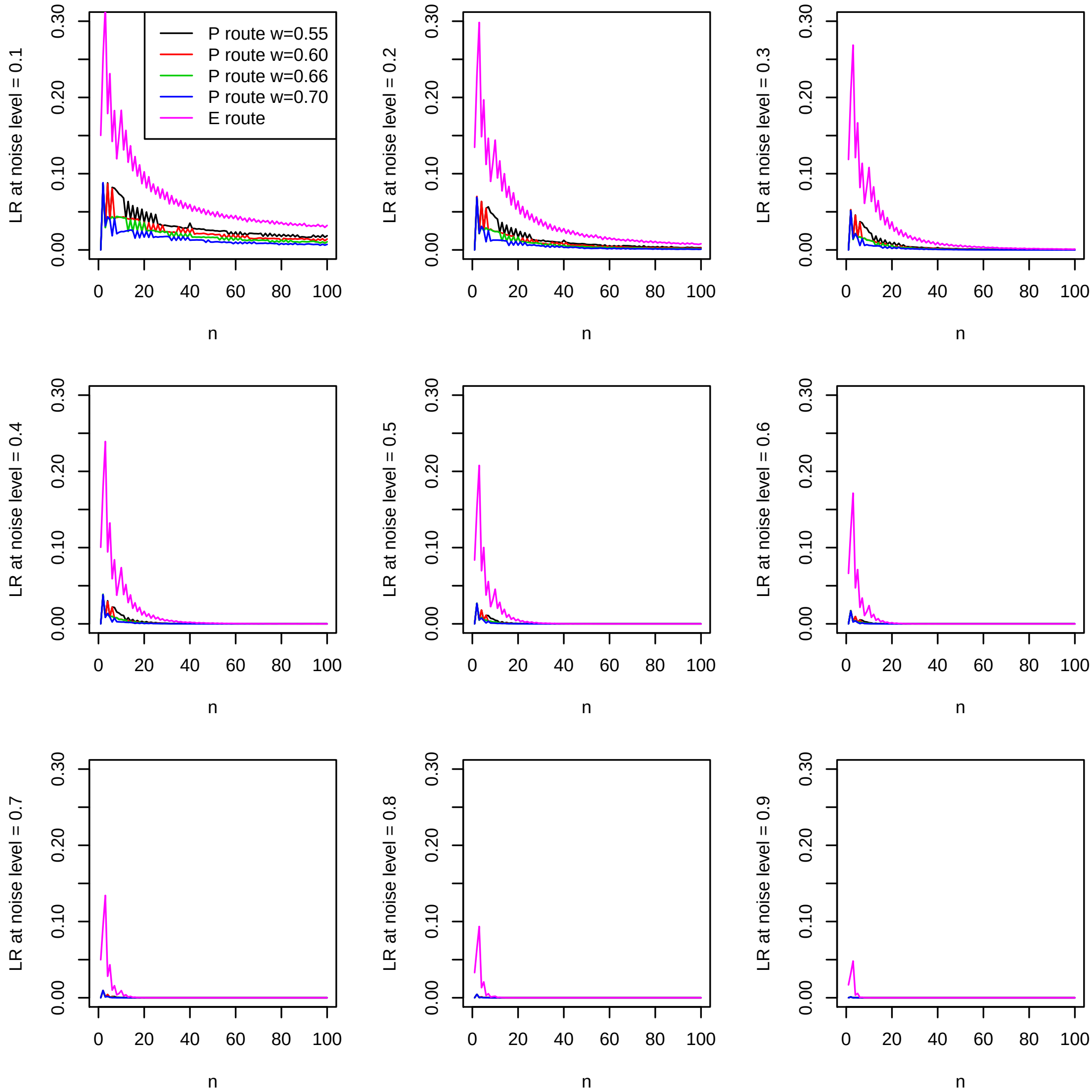

Figure 5.14: With average market condition, LR of both trading routes 2.1 and 2.2 under different noise levels p_2 for a call option expired data at $T = 100$.

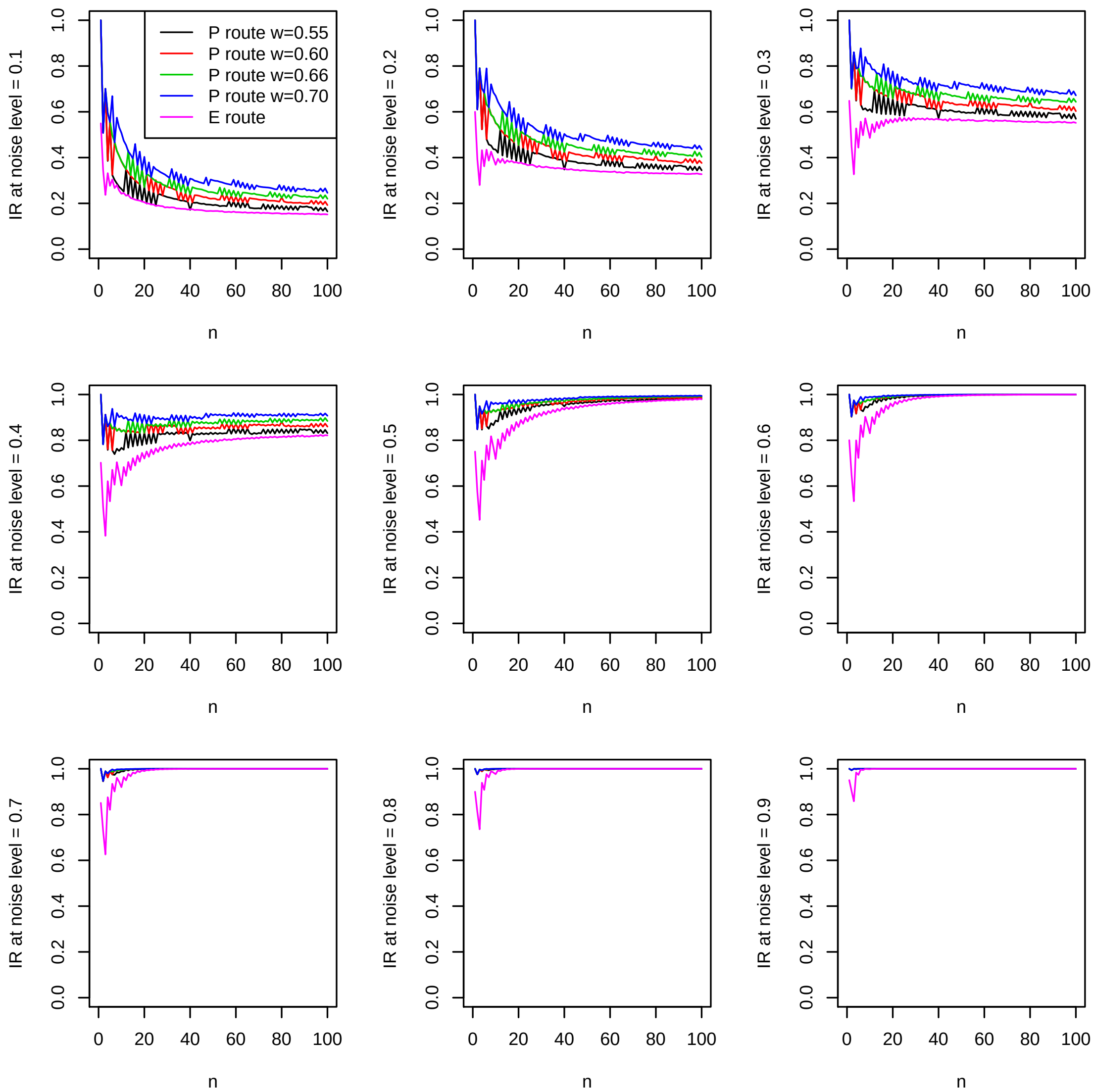

Figure 5.15: With average market condition, IR of both trading routes 2.1 and 2.2 under different noise levels p_2 for a call option expired data at $T = 100$.

present value payoff in the long run. (See Figure 5.12)

Performance of NPI call option trading routes under different market conditions given imprecise data available

Subject to the different noise levels in data, the performance of NPI European call option trading routes under different market conditions are examined below.

It is observed from simulations that both NPI call option trading routes effectively and efficiently recognize the noise from the imprecise data and gradually take less trading action as the noise level increase. Since both NPI call trading routes have similar patterns of decaying phenomenon on the average present value surface $\overline{f_i^C}$, and one complete example require nine pages of space, for the sake of brevity, we only present one complete example average present value $\overline{f_i^C}$ surface for route 2.1 with threshold value $w = 0.6$ in the Appendix B. (See Figure B.1-B.9)

To demonstrate the difference of reaction between route 2.1 and route 2.2 under different noise levels in a specific market condition, the $\overline{f_i^C}$, R_{wr}^C, R_{lr}^C and R_{ir}^C profiles of trading routes 2.1 and 2.2 under market condition $p_1 = 0.9$ for a call option expired at $T = 100$ are presented in Figures 5.16-5.19 below.

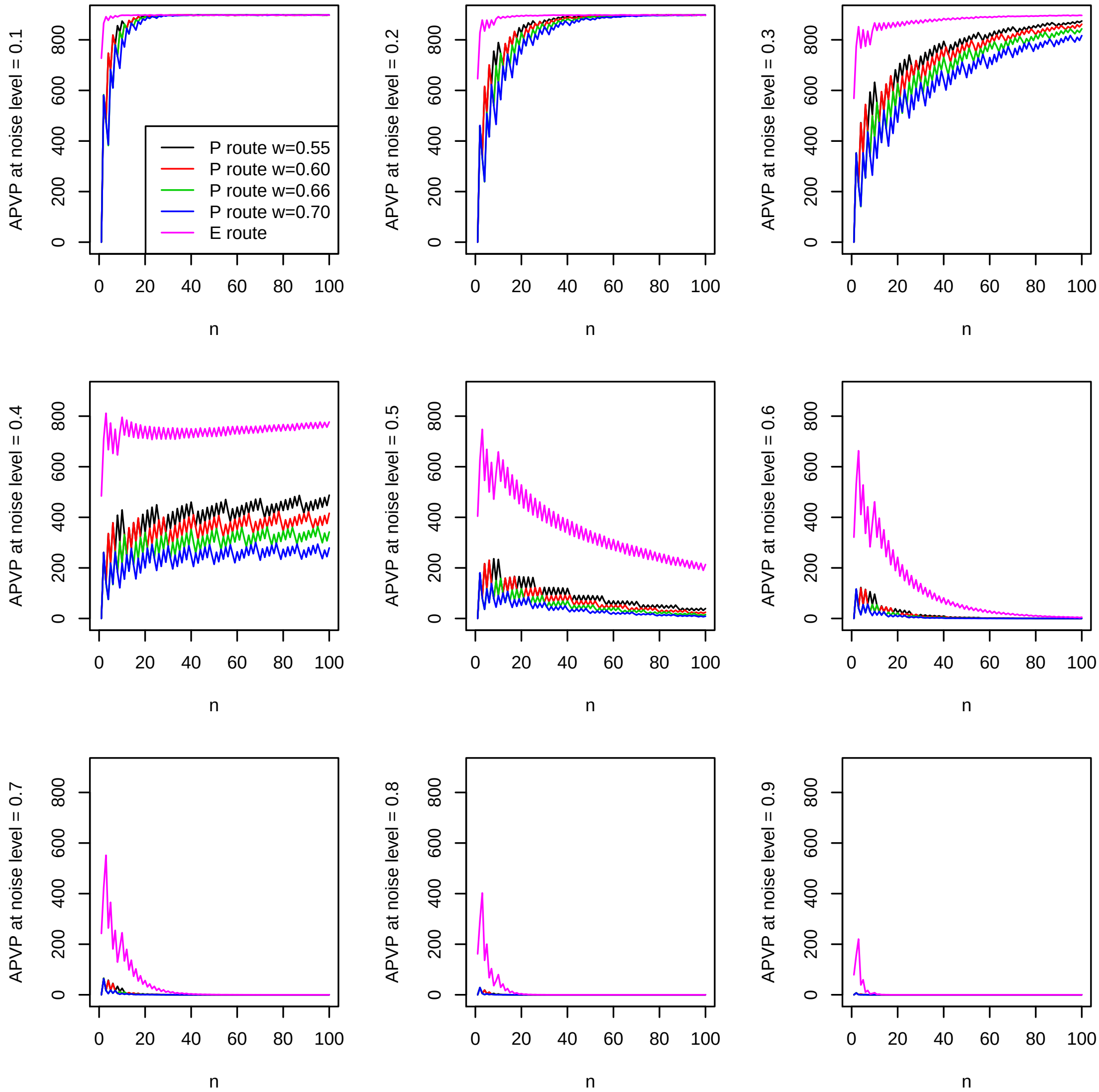

Figure 5.16: APVP of routes 2.1 and 2.2 for call option expired at $T = 100$ under market condition $p_1 = 0.9$ and different noise levels p_2.

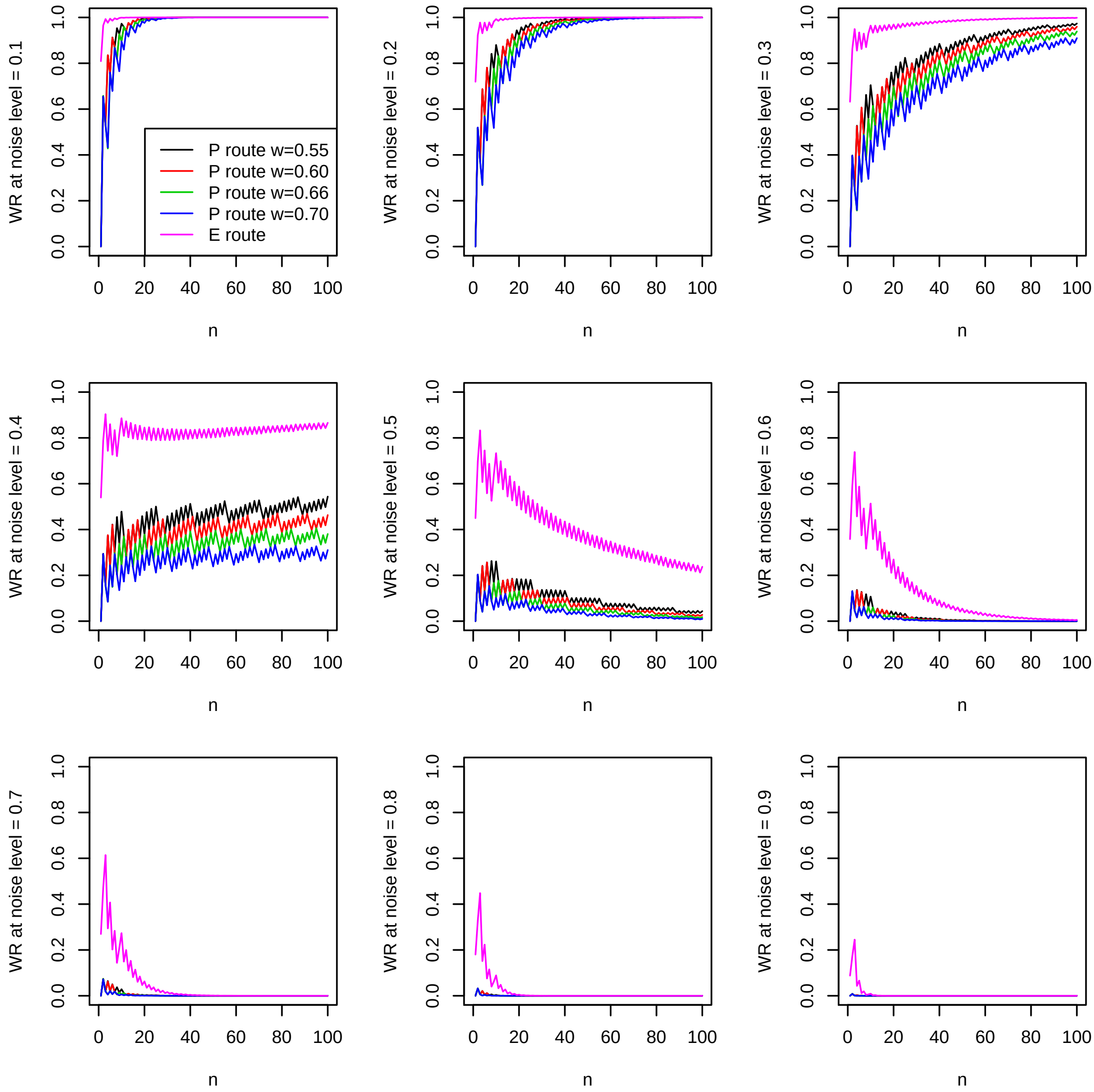

Figure 5.17: WR of routes 2.1 and 2.2 for call option expired at $T = 100$ under market condition $p_1 = 0.9$ and different noise levels p_2.

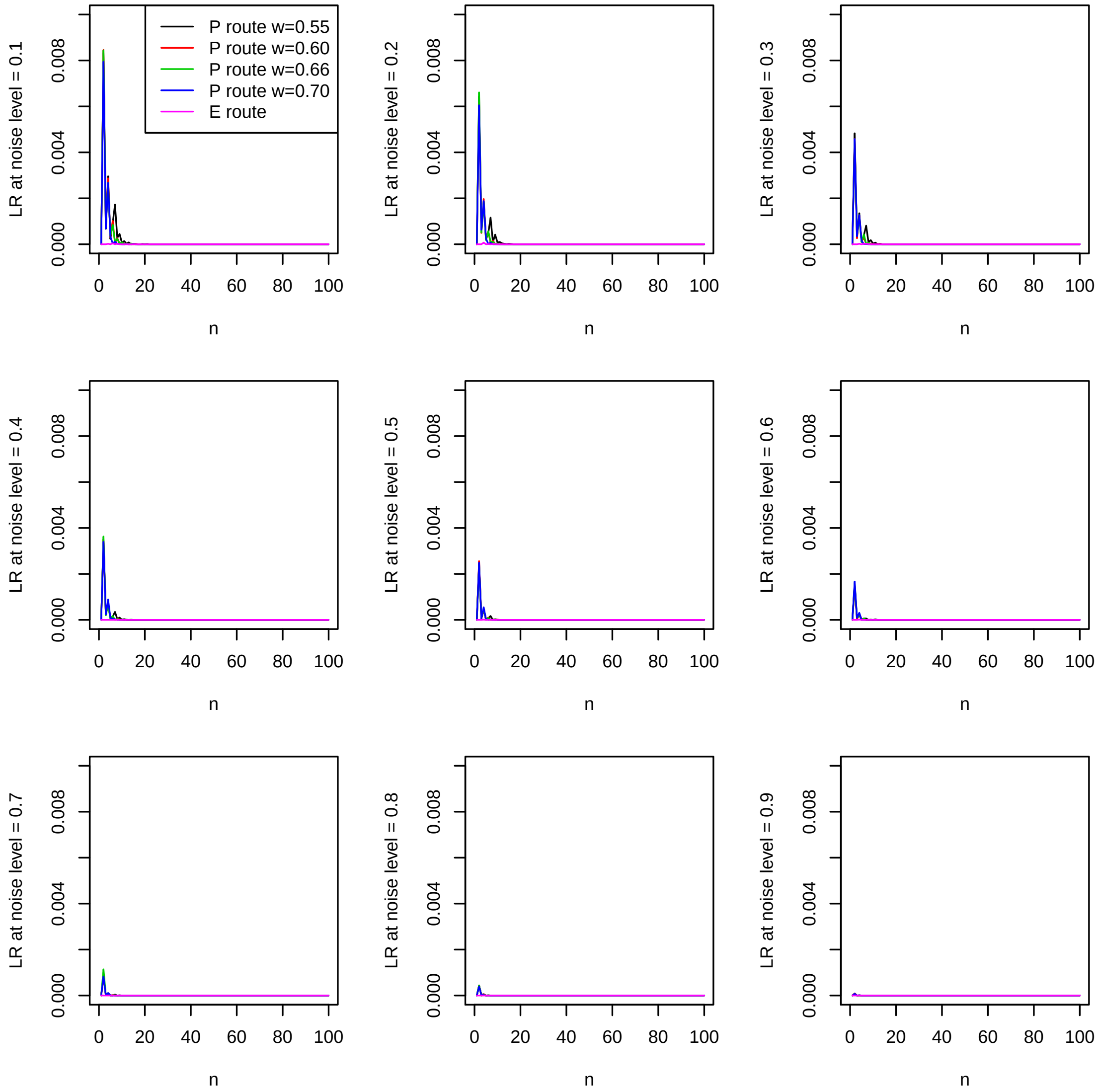

Figure 5.18: LR of routes 2.1 and 2.2 for call option expired at $T = 100$ under market condition $p_1 = 0.9$ and different noise levels p_2.

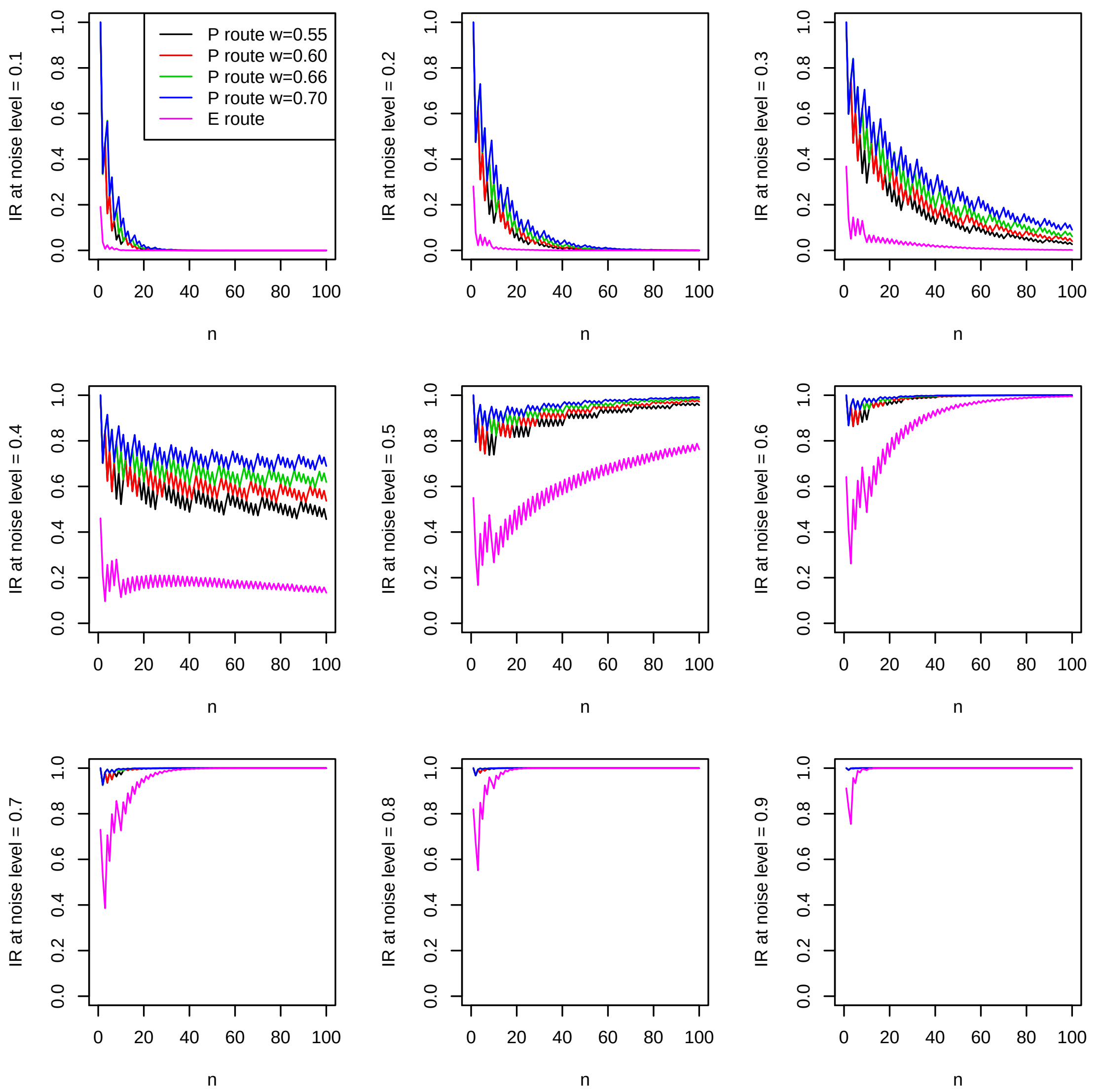

Figure 5.19: IR of routes 2.1 and 2.2 for call option expired at $T = 100$ under market condition $p_1 = 0.9$ and different noise levels p_2.

From Figure 5.19, it could found that route 2.1 is more sensitive to noise affection than route 2.2. More precisely, given the same amount of data available, route 2.1 as imprecise probability trading route has higher inaction than route 2.2 under different noise levels affection. Moreover, with a higher threshold value w, route 2.1 tends to be more conservative in trading. Although from Figure 5.18 under market condition $p_1 = 0.9$, both trading route 2.1 and route 2.2 seems to have similar loss rate. Simulation results from market condition $p_1 \in (0.1, 0.5)$ has shown that route 2.1 has a lower loss rate than route 2.2, especially when a higher threshold value is set. From Figure 5.17, it appears that the win rate of route 2.1 is significantly lower than route 2.2 when noise level $p_2 \in (0.4, 0.6)$. However, one should notice this is mainly attributed to higher inaction rate of route 2.1 when noise level $p_2 \in (0.4, 0.6)$ from Figure 5.19. In contrast, from Figure 5.16, it is found that route 2.2 as imprecise expectation trading route generally has higher average present value payoff than route 2.1 under different noise levels. Overall, one could conclude that both routes 2.1 and route 2.2 are able to recognize level under different market conditions and take less trading accordingly. Moreover, both trading routes maintain their primary trading objective under different market conditions and noise levels. With the same amount of noise level increase, route 2.1, as a more conservative trading route, has a higher increment in its inaction rate, resulting in quicker decay phenomenon on the $\overline{f_i^C}$ surface.

5.1.3 Overall review of European call option trading simulation

From Section 5.1.2, one could reach following conclusion:

The proposed NPI trading route 2.1 and route 2.2 have good performance under all market conditions and different noise levels. It is confirmed that both proposed trading route has quick learning in data, noise recognition, market condition recognition and predictability in nature.

Both trading routes are able to extract correct underlying information from the data effectively and efficiently and take corresponding correct action under different market conditions. The data learning process also has moderate noise resistance when a low noise level is presented. When the data is affected by a high noise level, both trading routes are able to recognize and stop taking any non sensible action readily.

Under no noise or low noise condition, given relatively enough data, route 2.1 has good risk control while route 2.2 is able to achieve higher average present value payoff throughout all different market conditions.

5.2 NPI method in European put option trading

In this section, we apply NPI method in European put option trading. Firstly, a simple put option trading scenario is specified. Under this scenario, two corresponding NPI put option trading routes are proposed. Subsequently, simulations are conducted to evaluate the performance of the proposed European put option trading routes by five different performance indexes.

5.2.1 Put option trading Scenario setting

Consider the scenario: one is allowed to long or short the one unit of put option with strike price K and maturity date T at a price $\Lambda_p^{\mathbb{Q}}(a_0, K)$ in time 0. Also, one is allowed to invest or borrow $\Lambda_p^{\mathbb{Q}}(a_0, K)$ with risk free interest rate r at time 0. Whatever position one enters, one has to keep the position for time length T and is obligated to close all risk position at time T (One is allowed to buy, sell or short sell the asset at price A_T for closing the risk position in time T). How should one, who is a NPI imprecise probability believer, without using any of his or her capital, make one's decision in trading to maximize one's capital gain in present value probabilistically or expectationally at time T? (Assume one's capital is able to cover any potential loss)?

In the scenario, the key points to emphasize are: fixed entering position time point, fixed closing position time point, one single put is available to long or short.

One is interested in the put option payoff $\Lambda_p(A_T, K) = (K - A_T(S_T))^+$ at time T. Since $\Lambda_p(A_T, K) = (K - A_T(S_T))^+$ is monotonically decreasing function of A_T and $A_T(\cdot)$ is monotonically increasing function S_T. Thus $\Lambda_p(A_T, K)$ is monotonically increasing function of S_T. One therefore can compute $\underline{E}(\Lambda_p(A_T, K))$,$\overline{E}(\Lambda_p(A_T, K))$ by construct $\underline{p}_{\Lambda_p(A_T,K)}(\cdot)$,and $\overline{p}_{\Lambda_p(A_T,K)}(\cdot)$ using Formulas 2.2.31 and 2.2.32.

A NPI believer, who prefer to use imprecise probability operator $\underline{p}$ and $\overline{p}$ could use following on European put option trading route in this scenario at time 0:

Set threshold value $0.5 < w < 1$. From NPI setting, one has $0 < \underline{p}((K - A_T(S_T))^+ > B(T)\Lambda_p^{\mathbb{Q}}(a_0, K)) < \overline{p}((K - A_T(S_T))^+ > B(T)\Lambda_p^{\mathbb{Q}}(a_0, K)) < 1$ if $S_T \subsetneq \mathbb{N}_0^T$ and $S_T \neq \emptyset$.

Imprecise probability European put trading route 2.3:

$$
\begin{cases}
\text{Borrow cash } \Lambda_p^{\mathbb{Q}}(a_0, K) \text{ and buy the put option at time 0,} \\
\text{Buy the asset at time } T \text{ and immediately exercise the option to sell the asset with} \\
\text{price K at time } T \text{ if } K > A_T \text{, return cash } \Lambda_p^{\mathbb{Q}}(a_0, K)B(T) \text{ to the lender at time } T \\
\text{if } \underline{p}[(K - A_T(S_T))^+ > B(T)\Lambda_p^{\mathbb{Q}}(a_0, K)] > w \\
\\
\text{Short the put option at time 0 for } \Lambda_p^{\mathbb{Q}}(a_0, K) \text{, invest } \Lambda_p^{\mathbb{Q}}(a_0, K) \text{ for risk free rate } r \\
\text{close all the position at time } T \text{ if } \underline{p}[(K - A_T(S_T))^+ < B(T)\Lambda_p^{\mathbb{Q}}(a_0, K)] > w \\
\\
\text{No action if none of above satisfied}
\end{cases}
$$

Motivation behind NPI imprecise probability European put option trading route 2.3:

Consider the event $(K - A_T(S_T))^+ > B(T)\Lambda_p^{\mathbb{Q}}(a_0, K)$ that the payoff of the put option at time T is greater than the interest $B(T)\Lambda_p^{\mathbb{Q}}(a_0, K)$ generated at time T by $\Lambda_p^{\mathbb{Q}}(a_0, K)$ at risk free rate r at time 0. If the lower probability of this event is greater than threshold value w ($w > 0.5$), then one would prefer to buy this put option and expect to earn more than $B(T)\Lambda_p^{\mathbb{Q}}(a_0, K)$ from the put option payoff in future time T.

On the contrary, consider the event $(K - A_T(S_T))^+ < B(T)\Lambda_p^{\mathbb{Q}}(a_0, K)$ that the payoff the put option at future time T is less than interest generated by investing $\Lambda_p^{\mathbb{Q}}(a_0, K)$ at time 0. If the lower probability of this event is greater than threshold value w ($w > 0.5$), then one would expect the payoff of put option more likely to be less than the interest generated by investing $\Lambda_p^{\mathbb{Q}}(a_0, K)$ at time 0. Thus one would prefer to short sell the put option and invest the amount $\Lambda_p^{\mathbb{Q}}(a_0, K)$ with risk free rate r.

If none of above conditions are satisfied, one would better off take no action as the lower probability of the desirable event is not significant enough for one to make a confident decision.

One can show that only one of actions could be taken in Route 2.3: Using inequality and conjugacy property of imprecise probability, one could know:

if $\underline{p}[(K - A_T(S_T))^+ > B(T)\Lambda_p^{\mathbb{Q}}(a_0, K)] > w$

then, by congugacy property

$1 - \overline{p}[(K - A_T(S_T))^+ < B(T)\Lambda_p^{\mathbb{Q}}(a_0, K)] > w$

by imprecise probability inequality

$\underline{p}[(K - A_T(S_T))^+ < B(T)\Lambda_p^{\mathbb{Q}}(a_0, K)] < \overline{p}[(K - A_T(S_T))^+ < B(T)\Lambda_p^{\mathbb{Q}}(a_0, K)] < 1 - w < w$

Thus, only one action could be taken in the imprecise probability European put option trading route.

Under the presetting scenario, a NPI believer, who prefer to use imprecise expectation operator $\underline{E}$ and $\overline{E}$ could use following European put trading route at time 0:

Imprecise expectation European put trading route 2.4:

$$\begin{cases} \text{Borrow cash } \Lambda_p^{\mathbb{Q}}(a_0, K) \text{ and buy the put option at time 0,} \\ \text{at time } T \text{ buy the asset and immediately exercise the option to sell the asset} \\ \text{with price K if } K > A_T\text{, also return cash } \Lambda_p^{\mathbb{Q}}(a_0, K)B(T) \text{ to the lender} \\ \text{if: } \underline{E}[(K - A_T)^+] > B(T)\Lambda_p^{\mathbb{Q}}(a_0, K) \\ \\ \text{Short the put option at time 0 for } \Lambda_p^{\mathbb{Q}}(a_0, K)\text{, invest } \Lambda_p^{\mathbb{Q}}(a_0, K) \text{ for risk free rate } r \\ \text{close all the position at time } T \text{ if } \overline{E}[(K - A_T)^+] < B(T)\Lambda_p^{\mathbb{Q}}(a_0, K) \\ \\ \text{No action if none of above satisfied} \end{cases}$$

Motivation behind NPI imprecise expectation European put option trading route 2.4:

When the lower expectation of put option payoff $\underline{E}[(K - A_T)^+]$ at future time T is greater than the interest $B(T)\Lambda_p^{\mathbb{Q}}(a_0, K)$ generated at time T by borrowing

$\Lambda_p^{\mathbb{Q}}(a_0, K)$ at time 0 into risk free interest rate r, one would prefer to borrow $\Lambda_p^{\mathbb{Q}}(a_0, K)$ at time 0 and buy the put option and expect to receive at least the amount of $\underline{E}[(K - A_T)^+] - B(T)\Lambda_p^{\mathbb{Q}}(a_0, K)$ at time T.

If the upper expectation of put option payoff $\overline{E}[(K - A_T)^+]$ at future time T is less than interest $B(T)\Lambda_p^{\mathbb{Q}}(a_0, K)$ generated at time T by investing $\Lambda_p^{\mathbb{Q}}(a_0, K)$ at time 0 into risk free interest rate r, one would rationally choose to short the put option and invest the money received into risk free rate, expecting monetary value in hand at time T is at least $B(T)\Lambda_p^{\mathbb{Q}}(a_0, K) - \overline{E}[(K - A_T)^+]$ by closing all risk position.

It is also easy to show that only one of actions could be taken in Route 2.4: In the imprecise probability framework, one has $\underline{E}[(K - A_T)^+] \leq \overline{E}[(K - A_T)^+]$ and therefore $\underline{E}[(K - A_T)^+] > B(T)\Lambda_p^{\mathbb{Q}}(a_0, K)$ and $\overline{E}[(K - A_T)^+] < B(T)\Lambda_p^{\mathbb{Q}}(a_0, K)$ could not be satisfied at the same time.

5.2.2 Simulation of put option trading in NPI Bernoulli model

In this section, we use simulation to study the performance of two proposed NPI European put option trading routes in the prescribed scenario setting.

We only present simulation results with following valued predefined parameters r, u, d, a_0 and K. Other value of predefined parameters values are also simulated, they all have similar patterns.

[Predefined parameters value r, u, d, a_0 and K] We use the same predefined parameter value for r, u, d, a_0 as asset trading chapter and put option strike price K is set at $K = 98$

All the decisions routes are simulated 100,000 times using the statistical software R version 3.5.1. The data generating process of underlying asset price is the same as in the asset trading simulation which will not be repeatedly stated here.

Performance evaluation function f_i^P

The performances of NPI European put trading routes are measured by five statistics of the present value pay-off function $f_i^P(n,T,i)$ in 100000 simulations. $f_i^P(n,T,i)$ is defined as follow:

$$f_i^P(n,T,i) = \begin{cases} (K-A_T(S_T))^+B(T)^{-1} - \Lambda_p^{\mathbb{Q}}(a_0,K) & \text{if first action of the trading route is taken} \\ \Lambda_p^{\mathbb{Q}}(a_0,K) - (K-A_T(S_T))^+B(T)^{-1} & \text{if second action of the trading route is taken} \\ 0 \quad \text{if no action} \end{cases}$$

where the inputs:
n is the length of historical asset price data one could learn;
T is the future time that the this function is evaluate;
$i \in (1, 100000)$ is the index of that particular simulation trial.

Five performance statistics of this function measure from 100000 simulations are:

Average present value payoff $\overline{f_i^P} = \dfrac{\sum_i f_i^P}{100000}$ Win-loss ratio $R_{wl}^P = \dfrac{|\{i : f_i^P > 0\}|}{|\{i : f_i^P < 0\}|}$

Win rate $R_{wr}^P = \dfrac{|\{i : f_i^P > 0\}|}{100,000}$ Loss rate $R_{lr}^P = \dfrac{|\{i : f_i^P) < 0\}|}{100,000}$

Inaction rate $R_{ir}^P = \dfrac{|\{i : f_i^P = 0\}|}{100,000}$

Sample simulation trials of different put option trading routes given precise or imprecise data

Several simulation trials are provided to illustrate how each put option trading routes work in the simulation process.

Simulation trial 1 Underlying market condition $p = 0.1$ (For the investor, this information is hidden), one observes following precise data of a asset in past 7 time stages

Time stage	-7	-6	-5	-4	-3	-2	-1
Data ($p = 0.1$)	0	0	0	0	0	0	0
Equivalently $(n, j) = (7, 0)$							

With predefined parameters value, one needs to make a decision whether or not enter a risk position of a put option of which the mature time is at time $T = 7$. By CRR pricing model, $q = \frac{e^r - d}{u - d} = 0.5044$ and current market price $\Lambda_p^{\mathbb{Q}}(a_0, K)$ is

$$\begin{aligned}
\Lambda_p^{\mathbb{Q}}(a_0, K) &= B(T)^{-1} \sum_{S_T=0}^{T} \binom{T}{S_T} (K - A_0 u^{S_T} d^{T-S_T})^+ q^{S_T} (1-q)^{T-S_T} \\
&= B(7)^{-1} \sum_{S_T=0}^{7} \binom{7}{S_T} (98 - 100 * 1.03^{S_T} (\frac{1}{1.03})^{7-S_T})^+ q^{S_T} (1-q)^{T-t-S_T} \\
&= 2.0094
\end{aligned}$$

If one uses route 2.3 (imprecise probability trading route) and set threshold value $w = 0.65$. One firstly finds out m such that $(K - A_7(m))^+ = B(T)\Lambda_p^{\mathbb{Q}}(a_0, K)$. $m \approx 2.8060$, One then find out $m_1 = \lceil m \rceil = 3$, $m_2 = \lfloor m \rfloor = 2$ and calculate $\underline{p}_{(7,0)}(S_7 \leq m_2) = 0.9038 > w$ and $\underline{p}_{(7,0)}(S_7 \geq m_1) = 0 < w$, thus one will take first action of route 2.3.

If one uses route 2.4 (imprecise expectation trading route), one will find out $\underline{E}_{(7,0)}[(K - A_T)^+] = 12.4349 > \Lambda_p^{\mathbb{Q}}(a_0, K)B(7) = 2.0193$ and $\overline{E}_{(7,0)}(A_7) = 16.6909 > \Lambda_p^{\mathbb{Q}}(a_0, K)B(7) = 2.0193$. Thus, one will take first action route 2.4 and execute the same strategy as one uses route 2.3.

Simulation trial 2 Underlying market condition $p = 0.9$, one observes following data of a asset in past 7 time stages

Time stage	-7	-6	-5	-4	-3	-2	-1
Data ($p = 0.9$)	1	1	1	0	1	0	1
Equivalently $(n, j) = (7, 5)$							

One needs to decide whether or not enter a risk position of a put option which expired at future 7 time units. If one uses route 2.3 (imprecise probability trading route) and set threshold value $w = 0.6$. With the predefined parameters value, current market price $\Lambda_p^{\mathbb{Q}}(a_0, K)$ is still 2.0094. The value m such that $(K - A_7(m))^+ =$

$B(T)\Lambda_p^{\mathbb{Q}}(a_0, K)$ is still $m \approx 2.8060$. One then still find out $m_1 = \lceil m \rceil = 3$, $m_2 = \lfloor m \rfloor = 2$ calculate $\underline{p}_{(7,5)}(S_7 \leq m_2) = 0.0513 < w$ and $\underline{p}_{(7,5)}(S_7 \geq m_1) = 0.8569 > w$, therefore one uses route 2.3 will take the second action.

If one uses route 2.4 (imprecise expectation trading route), one will find out $\underline{E}_{(7,5)}[(K - A_T)^+] = 0.4951 < \Lambda_p^{\mathbb{Q}}(a_0, K)B(7) = 2.0193$ and $\overline{E}_{(7,5)}(A_7) = 1.3888 < \Lambda_p^{\mathbb{Q}}(a_0, K)B(7) = 2.0193$. Thus, one would take second action of route 2.4 in this trial.

Simulation trial 3 Underlying market condition $p_1 = 0.2$, and noise level $p_2 = 0.3$, one observes following data of a asset in past 7 time stages.

Time stage	-7	-6	-5	-4	-3	-2	-1
Data ($p_1 = 0.2$, $p_2 = 0.3$)	0	{0,1}	{0,1}	{0,1}	0	0	0
Equivalently $[n, J] = [7, (0, 3)]$							

One needs to make a decision as above. If one uses route 2.3 (imprecise probability trading route) and set threshold value $w = 0.65$. The value m such that $(K - A_7(m))^+ = B(T)\Lambda_p^{\mathbb{Q}}(a_0, K)$ is still $m \approx 2.8060$. Set $m_1 = \lceil m \rceil = 3$, $m_2 = \lfloor m \rfloor = 2$ and calculate $\underline{p}_{[7,(0,3)]}(S_7 \leq 2) = 0.2960 < w$ and $\underline{p}_{[7,(0,3)]}(S_7 \geq 3) = 0 < w$ so one will take no action in this case.

If one is using route 2.4 (imprecise expectation trading route), one will find out $\underline{E}_{[7,(0,3)]}[(A_T - K)^+] = 2.9773 > \Lambda_p^{\mathbb{Q}}(a_0, K)B(7) = 2.0193$ and $\overline{E}_{(7,(0,3))}(A_7) = 16.6909 > \Lambda_p^{\mathbb{Q}}(a_0, K)B(7) = 2.0193$. Thus, one will take first action of route 2.4 in this case.

Simulation trial 4 Underlying market condition $p_1 = 0.8$, and noise level $p_2 = 0.6$, one observes following data of a asset in past 7 time stages.

Time stage	-7	-6	-5	-4	-3	-2	-1
Data ($p_1 = 0.8$, $p_2 = 0.6$)	{0,1}	{0,1}	0	{0,1}	1	{0,1}	1
Equivalently $[n, J] = [7, (2, 6)]$							

With the same situation, one uses route 2.3 (imprecise probability route) and set threshold value $w = 0.6$. One calculate $\underline{p}_{[7,(2,6)]}(S_7 \leq 2) = 0.0105 < w$ and $\underline{p}_{[7,(2,6)]}(S_7 \geq 3) = 0.2797 < w$, so one will take no action in this case.

If one uses route 2.4 (imprecise expectation trading route), one will find out $\underline{E}_{[7,(2,6)]}[(K-A_T)^+] = 0.1041 < \Lambda_p^{\mathbb{Q}}(a_0, K)B(7) = 2.0193$ and $\overline{E}_{(7,(2,6))}[(K-A_T)^+] = 8.5749 > \Lambda_p^{\mathbb{Q}}(a_0, K)B(7) = 2.0193$. Thus, one who uses route 2.4 will take no action in this case.

Performance of NPI put option trading routes under average market condition given precise data available

Given precise data, under the average market condition, the performances of proposed NPI European put trading routes are evaluated and discussed below.

Given $n \in (1, 100)$ units of precise data, the average present value payoff of route 2.3 and route 2.4 for put option expired at future time $T \in (1, 100)$ is presented in Figure 5.20. It could be observed from Figure 5.20 that both trading routes 2.3 and route 2.4 produce positive average present value payoff for put options with different expiration dates. Both of trading routes have similar value in $\overline{f_i^P}$ for most combinations of the size of available data and expiration date of the put option. Specifically, one could observe with a small amount of data available, route 2.4 is able to generate higher $\overline{f_i^P}$ that route 2.3 due to the same reason explained in the call option trading section.

To examine the put option trading routes' performance in more detail, the performance indexes $\overline{f_i^P}$, R_{wl}^P, R_{wr}^P, R_{lr}^P and R_{ir}^P are presented in Figure 5.21

From Figure 5.21, it could be seen that both put option trading routes 2.3 and 2.4 exhibit similar features as routes 2.1 and 2.2 for call option trading. Both trading routes 2.3 and 2.4 has quick speed of learning in data. When 20 units of historical data becomes available, both trading routes are able to achieve close to its optimum performance in all indexes. As one may have expected, the expectation trading route 2.4 is able to produce higher $\overline{f_i^P}$, but the performance in risk control of R_{wl}^P and R_{lr}^P are worse than route 2.3. Nevertheless, it should not be neglected that both trading routes has loss rate less than 0.2 for put option expired in $T = 100$ when 15 units of data points become available. Also, both trading routes are able to avoid making non sensible trading when only a small amount of data is presented. (This is reflected in higher inaction rates for small n in Figure 5.21)

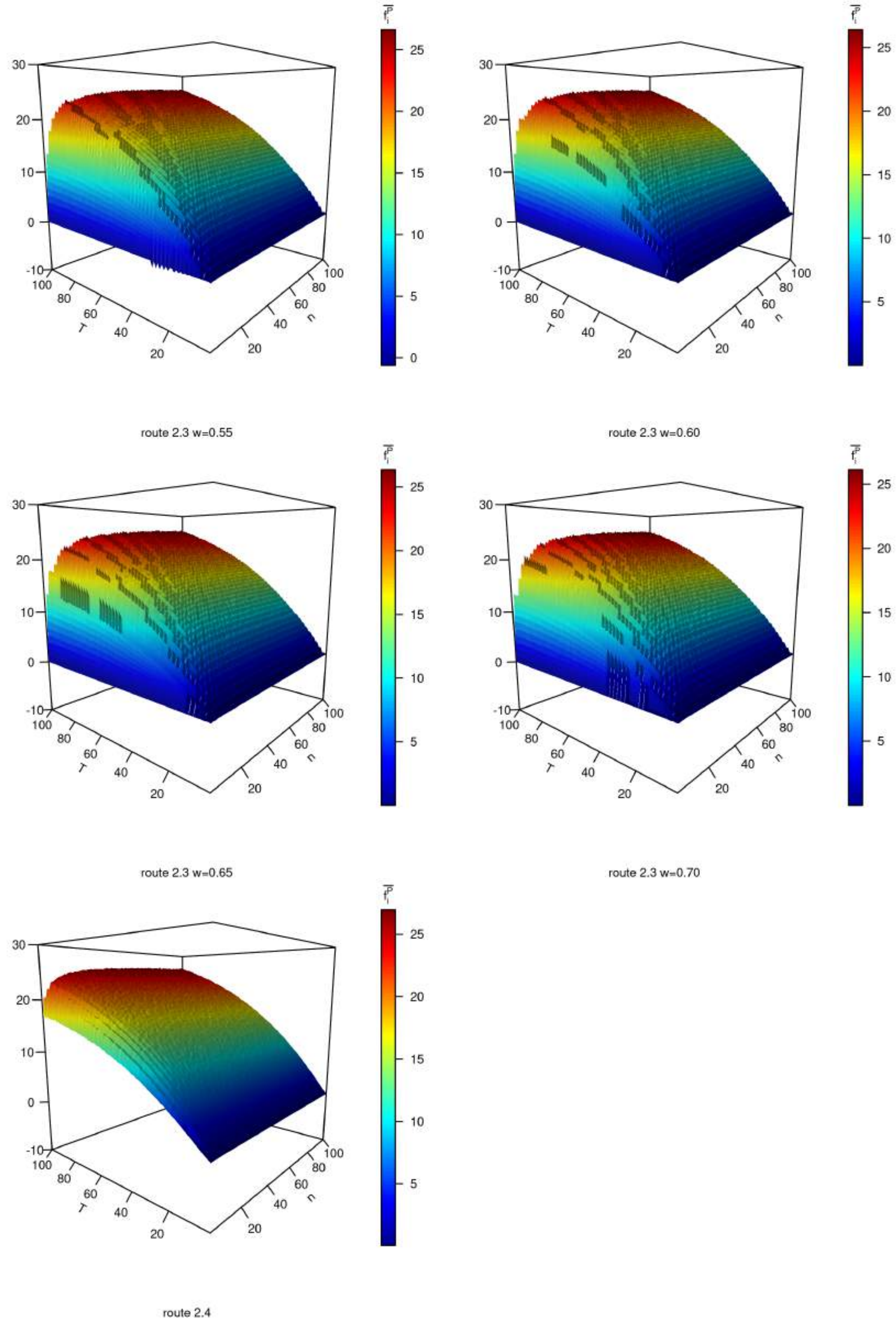

Figure 5.20: APVP of routes 2.3 and 2.4 for put option with different expiration date under average market condition given precise data.

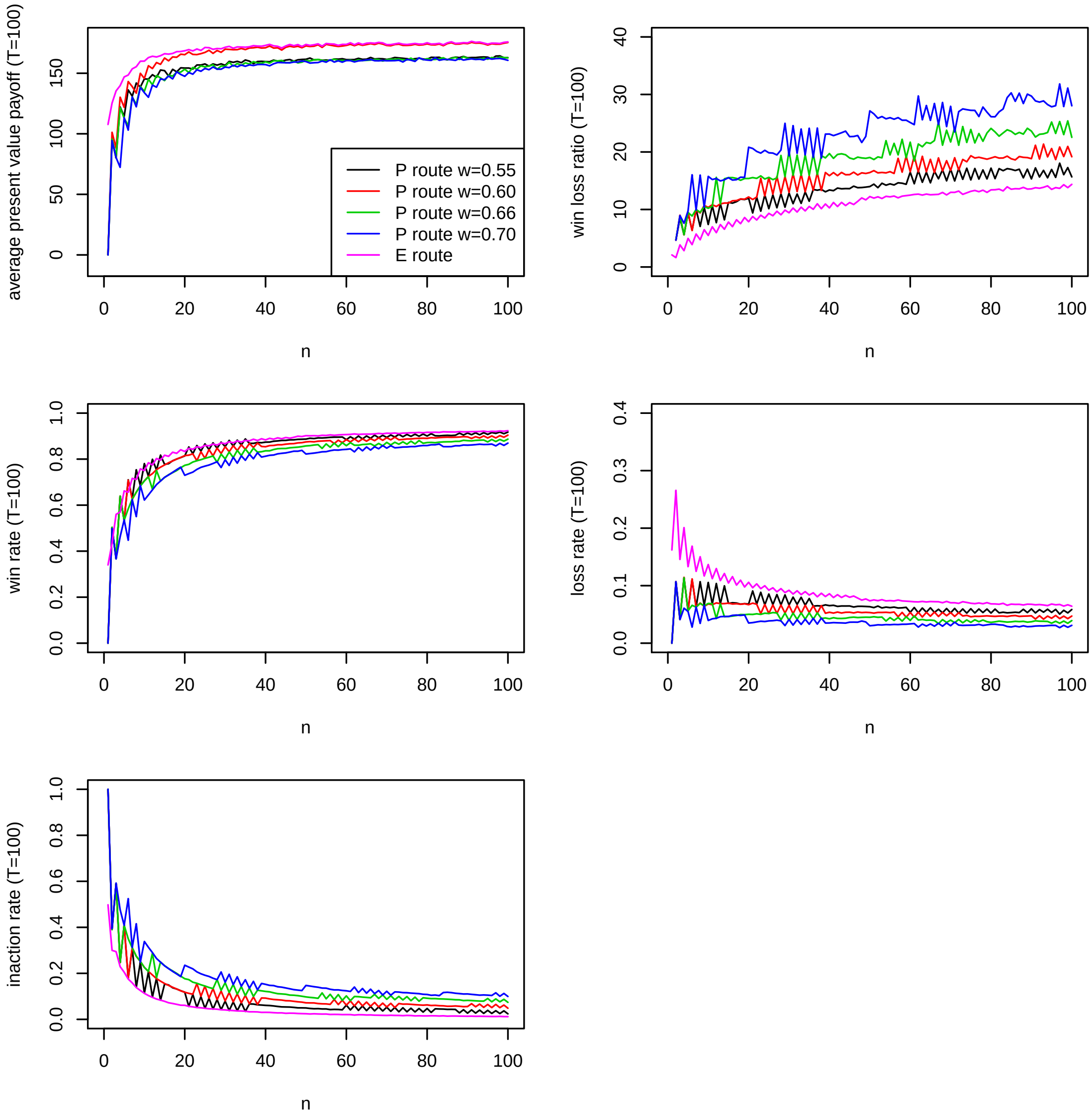

Figure 5.21: Performance comparison of routes 2.3 and 2.4 for put option expired in $T = 100$ under average market condition given precise data.

Performance of NPI put option trading routes under different market conditions given precise data available

Given precise data, the proposed NPI put option trading routes are further examined under different market condition below.

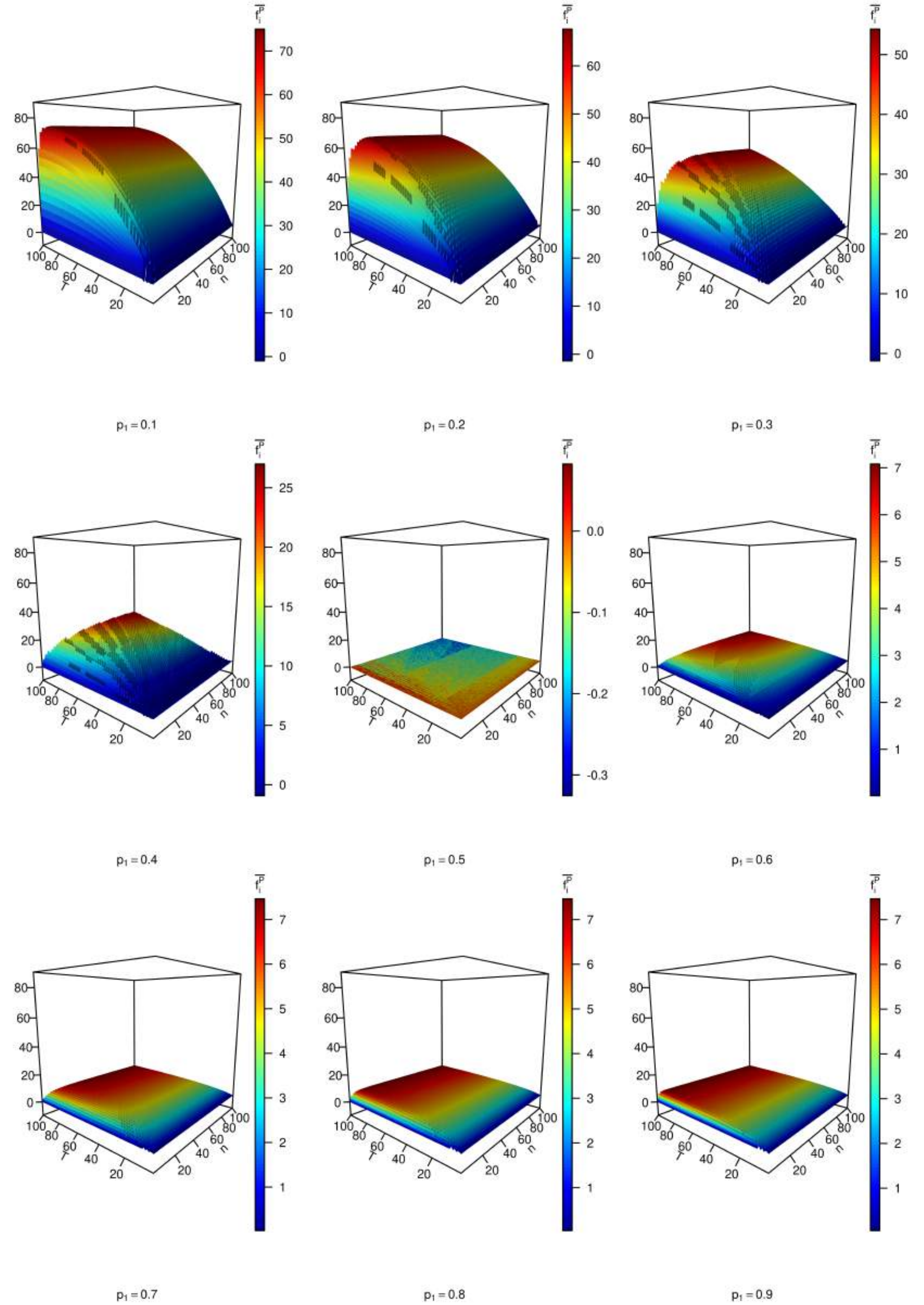

Figure 5.22: Under different market conditions, APVP of trading route 2.3 with threshold value $w = 0.6$.

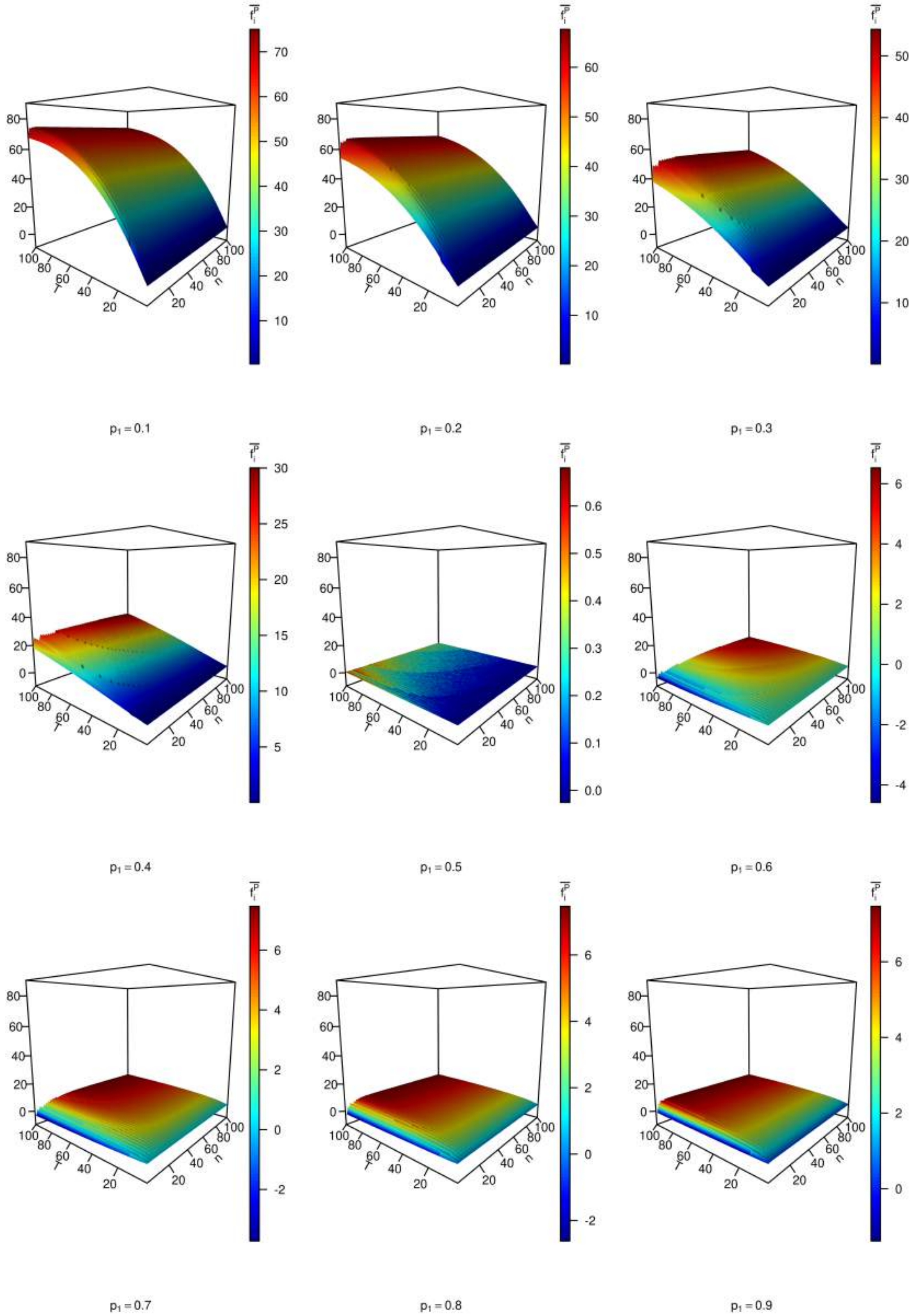

Figure 5.23: Under different market conditions, APVP of trading route 2.4.

Using the presetting parameter values in the simulation, one could know that for a put option which expired in time $T = 100$, the maximum present value payoff could achieved by action 1, namely borrowing cash and buying the put option, is $(K - a_0 d^{100})B(100)^{-1} = 86.52308$ and the maximum present value payoff could be achieved by action 2, namely short selling the put option and invest the cash in risk free rate is $B(T)\Lambda_p^{\mathbb{Q}}(a_0, K) = 7.9670$.

When the market condition is declining for the underlying asset price ($p < 0.5$), one would have better chance to make a profit if one chooses to borrow cash and buy and the put option. On the contrary, if the market condition is surging for the underlying asset price ($p > 0.5$), one would have better chance to make a profit if one short sell the put option and invest the received cash into risk free rate.

From Figure 5.22 and Figure 5.23, one could see that when the more than 20 units of data point are gathered, both trading routes 2.3 and 2.4 are able to learn and recognize the market condition from the data, also execute the correct action according to the market condition (As described above). Although average trading loss does happen in route 2.3 when the market condition is $p = 0.5$ and route 2.4 when the market condition is $p \in (0.6, 0.8)$, one should notice that the average trading loss only happens in the case where a small amount of data is available ($n < 5$) and when it happens, the loss amount is well controlled at a low level.

To have a better understanding of both trading routes, the performance indexes R_{wl}^P, R_{wr}^P, R_{lr}^P and R_{ir}^P of routes 2.3 and 2.4 for a put option expired at time $T = 100$ are plotted in Figures 5.24-5.28.

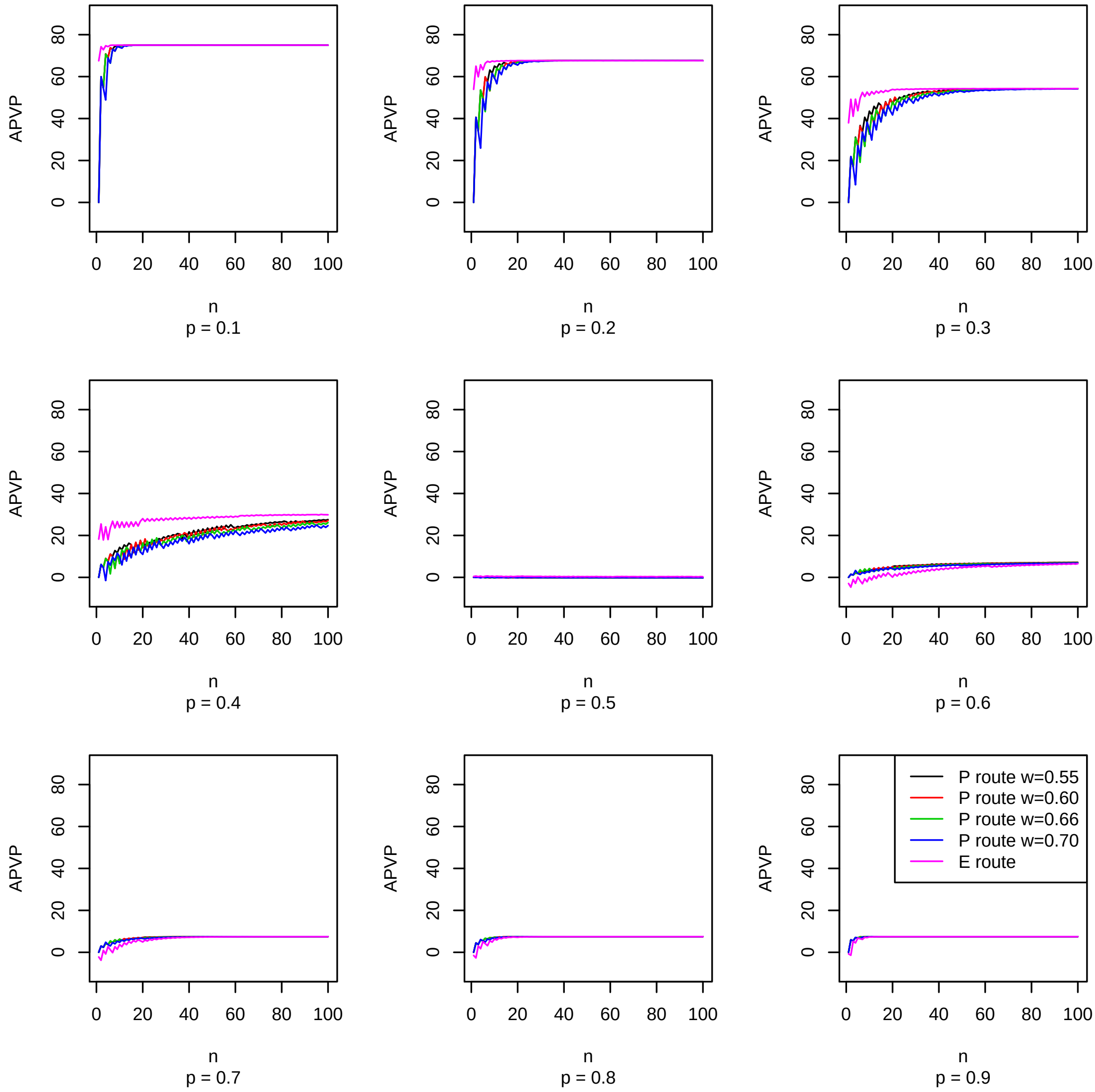

Figure 5.24: Under different market conditions, APVP of trading routes 2.3 and 2.4.

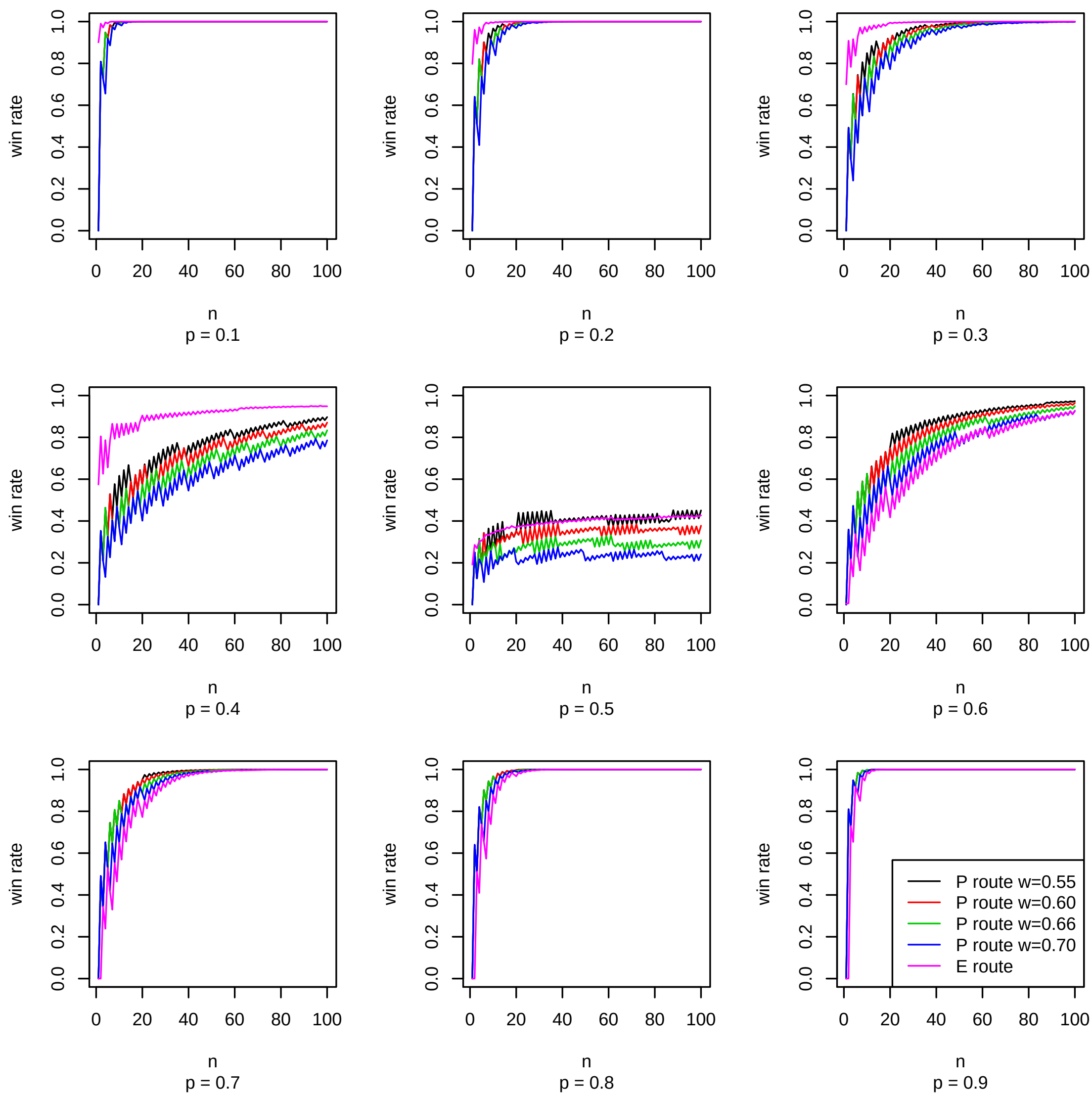

Figure 5.25: Under different market conditions, WR of trading routes 2.3 and 2.4.

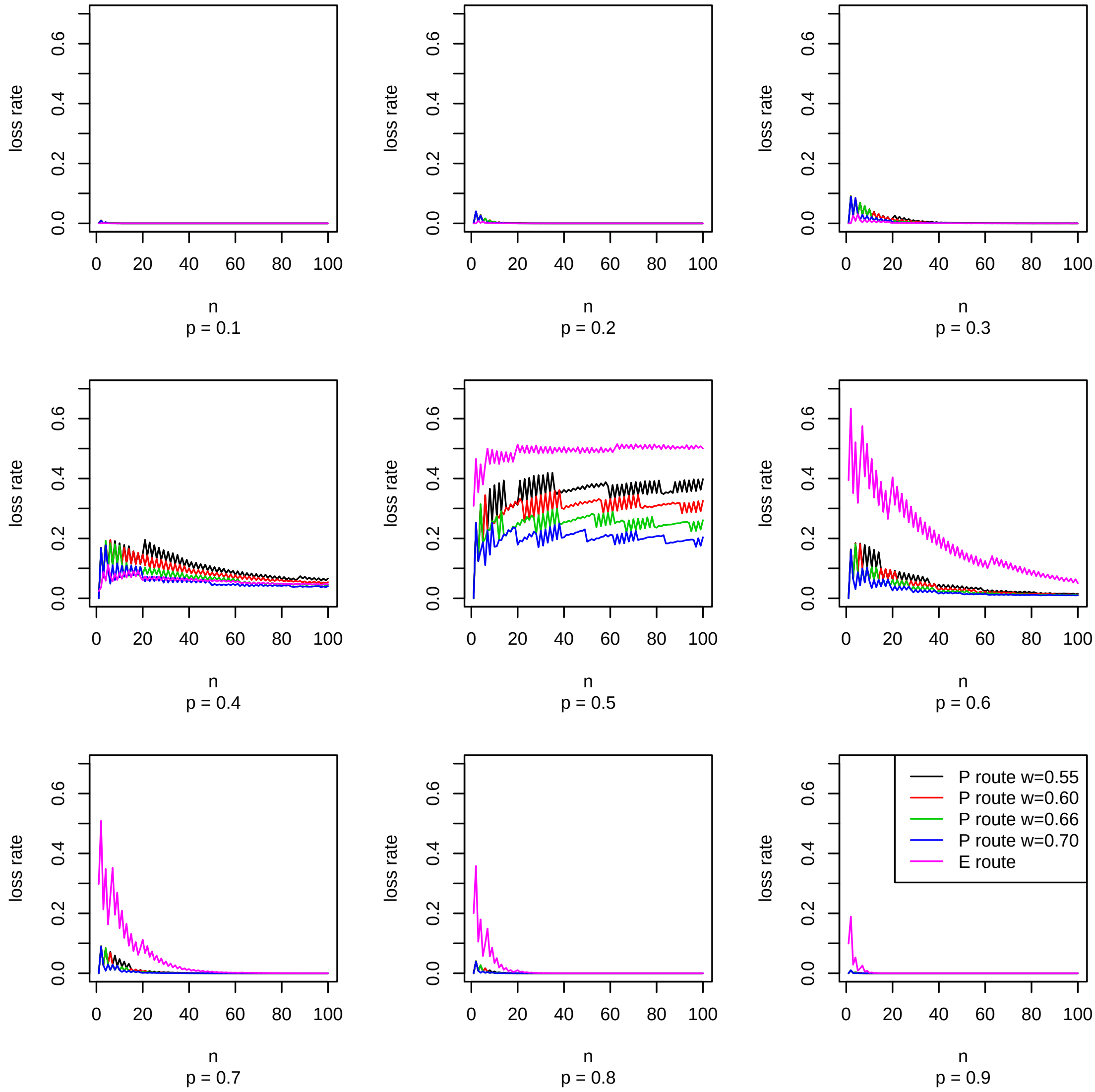

Figure 5.26: Under different market conditions, LR of trading routes 2.3 and 2.4.

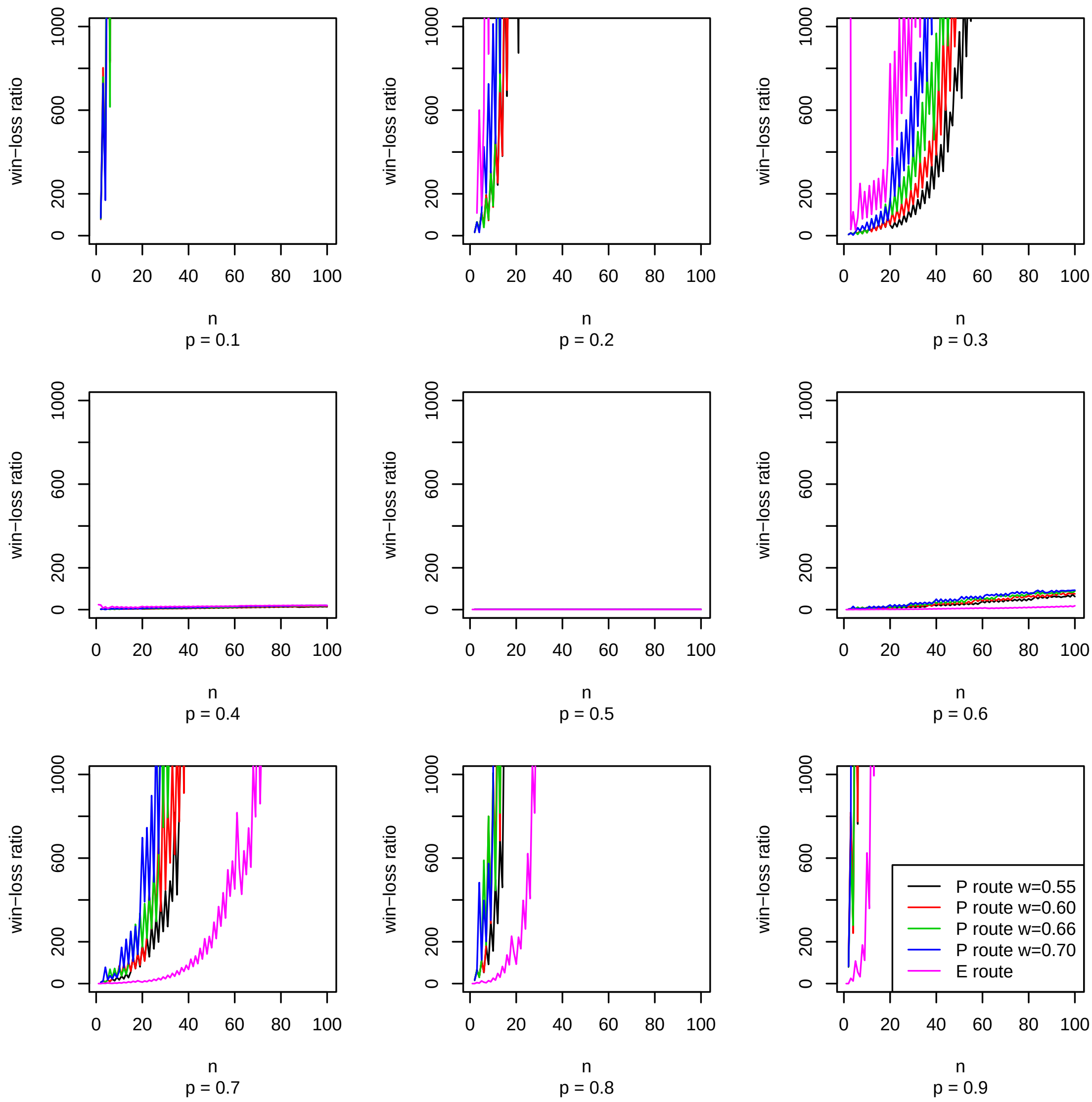

Figure 5.27: Under different market conditions, WLR of trading routes 2.3 and 2.4.

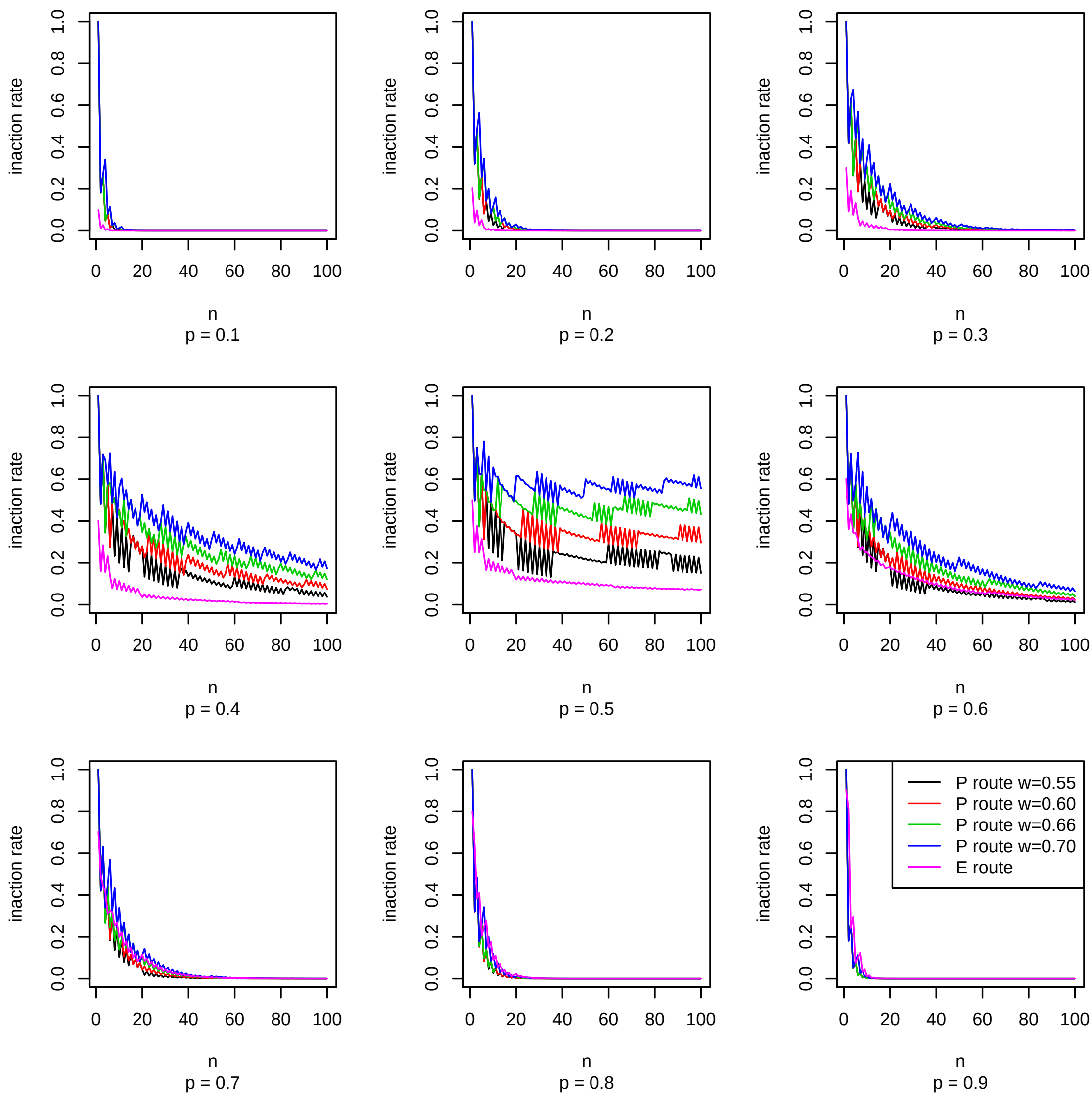

Figure 5.28: Under different market conditions, IR of trading routes 2.3 and 2.4.

From Figure 5.28, one could observe that with a small amount of data available, both routes 2.3 and 2.4 have extremely high inaction rate. In other words, both trading routes avoid taking non sensible action when there is only a limited amount of information is known. After enough data are gathered $n > 15$, both trading routes have higher inaction rate when the market condition is relatively neutral $p \in (0.4, 0.6)$. This is reasonable because in those conditions, neither taking action 1 nor action 2 would result in a positive payoff for a level of certainty. Moreover, one may notice route 2.4 is generally more active than route 2.3 throughout all market conditions, as its primary objective to achieve maximum payoff on average with less focused on the loss rate control.

From Figure 5.24, it could be seen that route 2.4 indeed has the primary objective to achieve maximum payoff on average. It has a higher payoff than route 2.3 when the market condition is favorable for buying the put option ($p < 0.5$). When the market condition is more favorable to short sell the put option ($p > 0.5$), both trading routes 2.3 and 2.4 have similar payoffs. Route 2.3, in contrast, is better at risk control. This could be confirmed from Figure 5.26. It overall has a lower loss rate than route 2.4 for all the market condition. Its threshold parameter w also offers one the flexibility for one to adjust this. The higher threshold parameter w one sets, the lower loss rate one could expect from route 2.3.

As regard to the win rate profile (See Figure 5.25), by staying more active, route 2.4 does have higher win rate when the market condition is favorable for buying the put option ($p < 0.5$). However, due to its lack of control in risk, its win rate is lower than route 2.3 when the market condition is more favorable for short sell the put option ($p > 0.5$). This results in higher win-loss ratio for route 2.4 when $p < 0.5$ and lower win-loss ratio for route 2.4 when $p > 0.5$. (See Figure 5.27)

Overall, it could be confirmed that both trading routes are able to learn and recognize the underlying market condition p from the data and execute the correct action accordingly. Route 2.3 is better at avoiding making losses in trading while route 2.4 is better at achieving higher present value payoff in different market conditions.

Performance of NPI put option trading routes under average market condition given imprecise data available

Under the average market condition, subject to the different noise levels, the performance of NPI European put option trading routes are evaluated below.

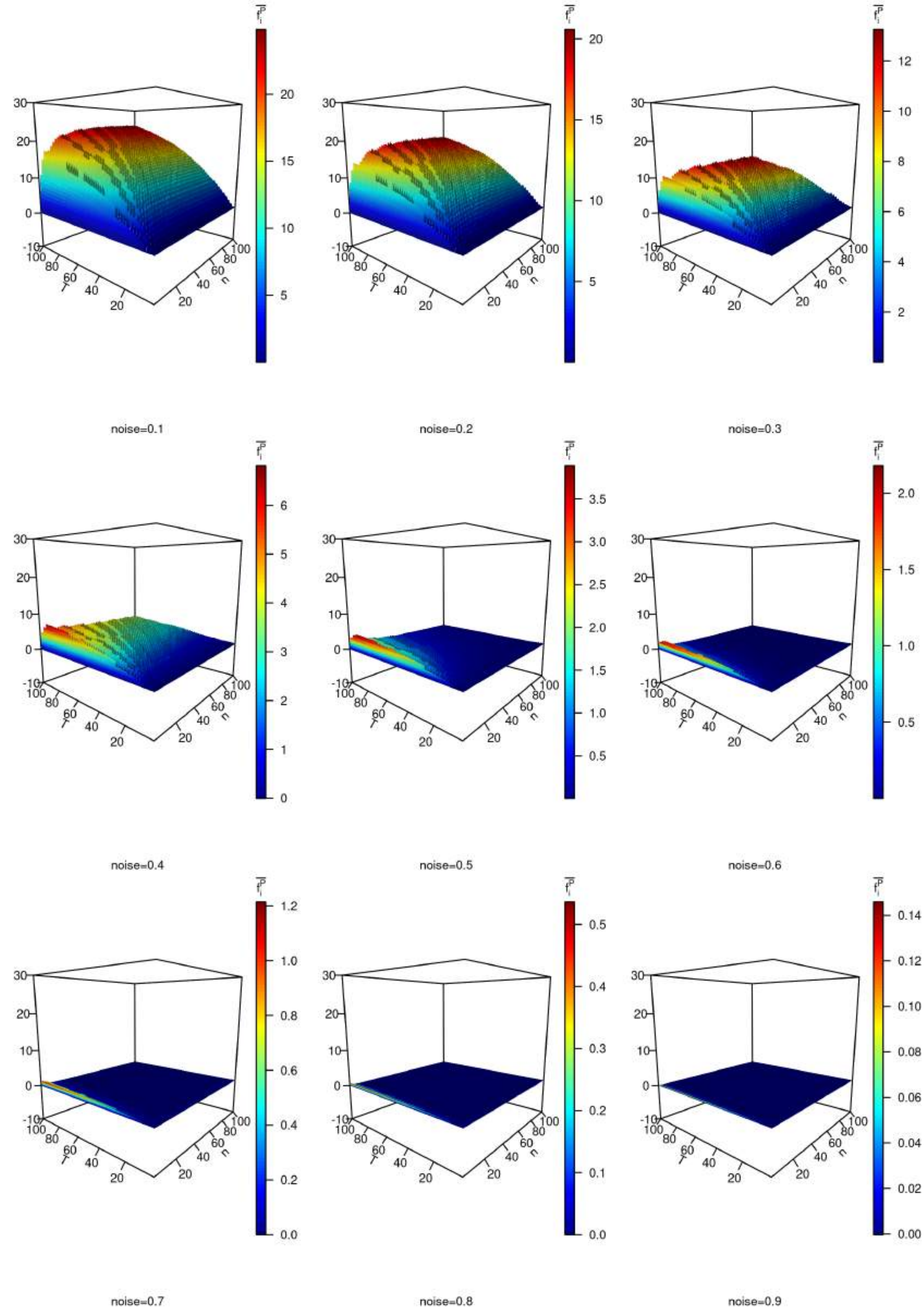

Figure 5.29: With average market condition, APVP of trading route 2.3 with threshold value $w = 0.6$ under different noise levels p_2.

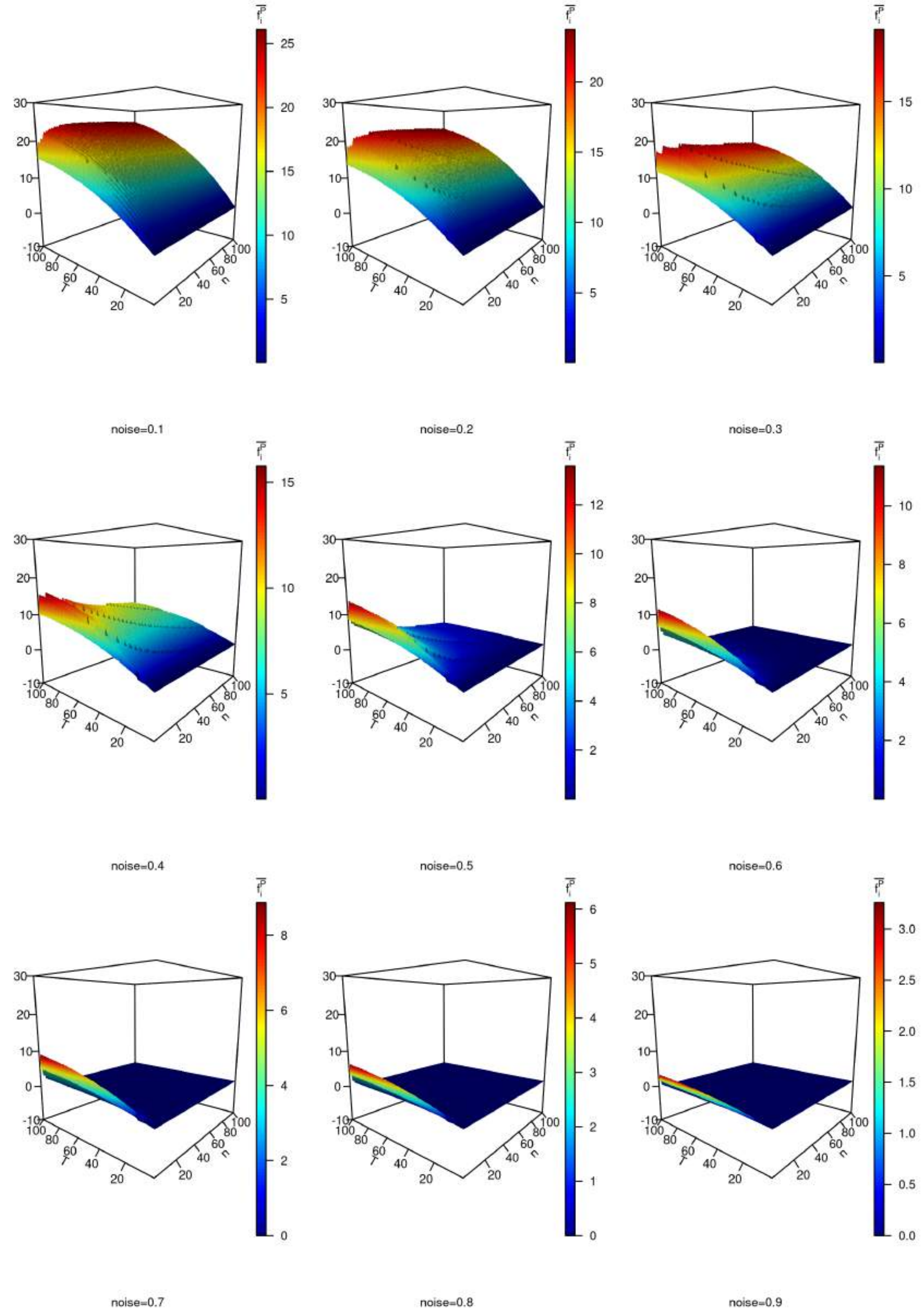

Figure 5.30: With average market condition, APVP of trading route 2.4 under different noise levels p_2.

Given n units data point, the average present value payoff $\overline{f_i^P}$ surfaces of route

2.3 with threshold value $w = 0.6$ and route 2.4 for different put option expired in 1 to 100 units time are plotted in Figure 5.29 and Figure 5.30 respectively.

When the data is affected by low level noise ($p_2 \leq 0.3$), both trading routes 2.3 and route 2.4 are still able to extract useful information of the market condition in each simulation trial and execute correct action accordingly. This results in similar patterns of $\overline{f_i^P}$ surfaces to the case where the data has no noise in the previous discussion.

With noise level p_2 increase, one could observe that the $\overline{f_i^P}$ surfaces of both trading routes 2.3 and 2.4 are decaying as more inactions are taking place in both trading routes. This indicates that both trading routes are able to recognize the noise level from the data. Also, with the same amount of noise level increase, route 2.3 stay more frequently inactive than route 2.4 due to its risk control nature. This can be seen from the value of z-axis that route 2.3 is decaying faster than route 2.4.

Overall, under the average market condition, both trading routes 2.3 and 2.4 maintain positive value in average present value payoff throughout all different noise levels.

Figures 5.31-5.34 present a more direct performance comparison of trading route 2.3 and route 2.4 by plotting the indexes $\overline{f_i^C}$, R_{wr}^C, R_{lr}^C and R_{ir}^C for a put option expired at time $T = 100$. As the noise level increases, the inaction rate of both trading routes 2.3 and 2.4 increase which results in close to zero loss rate. Therefore, win-loss ratio R_{wr}^P is not presented here.

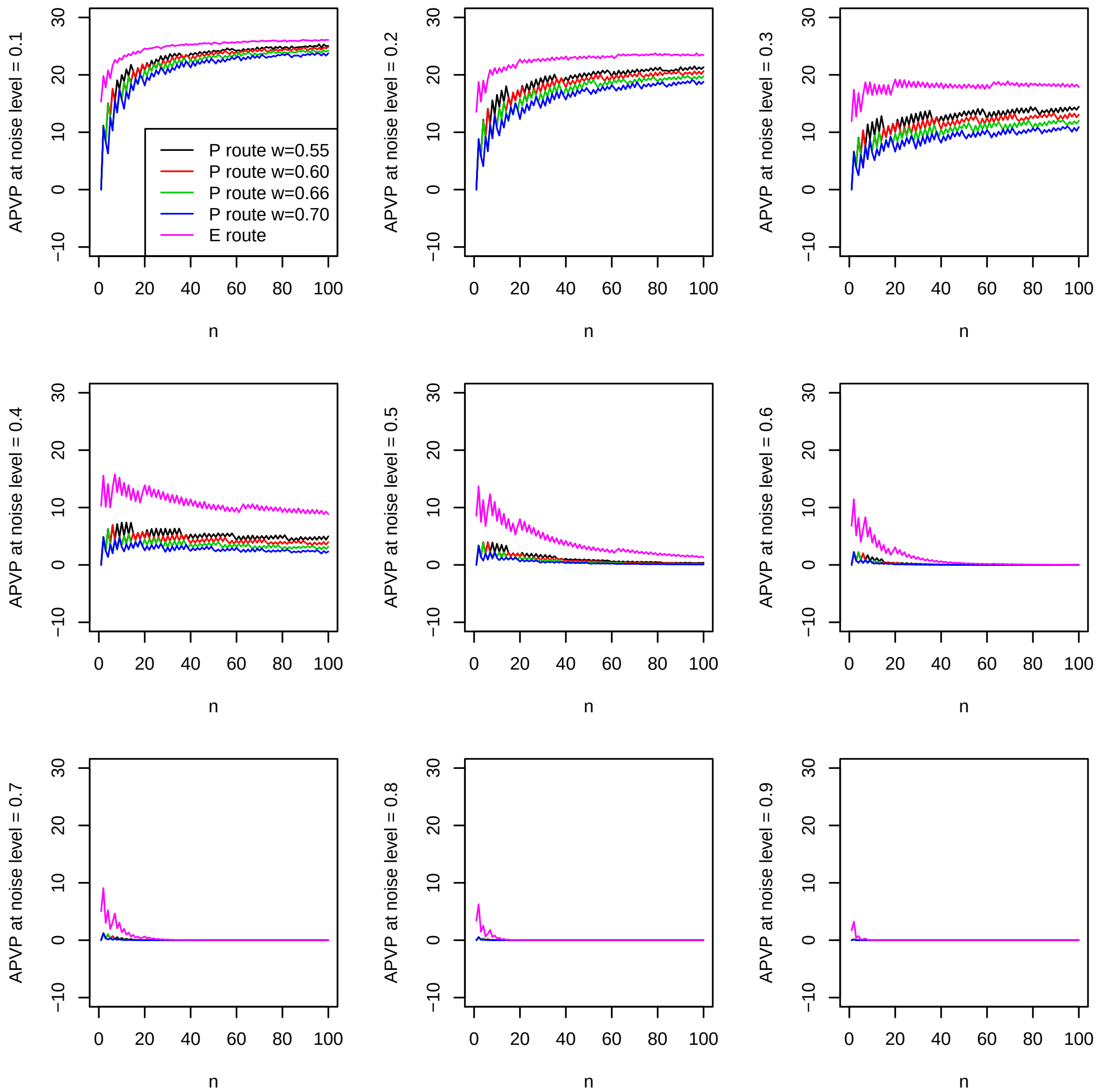

Figure 5.31: With average market condition, APVP of both trading routes 2.3 and 2.4 under different noise levels p_2 for a put option expired data at $T = 100$.

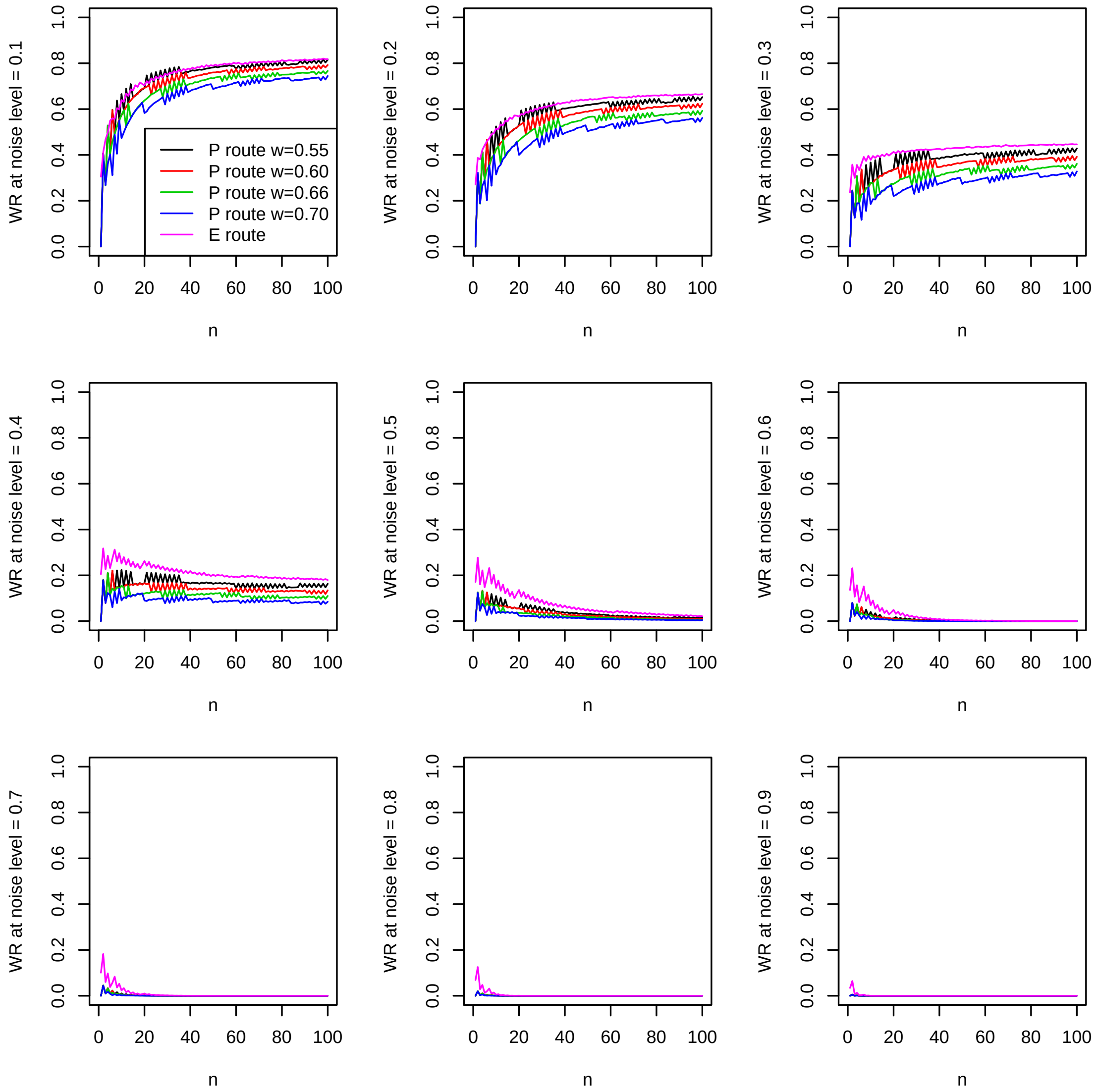

Figure 5.32: With average market condition, WR of both trading routes 2.3 and 2.4 under different noise levels p_2 for a put option expired data at $T = 100$.

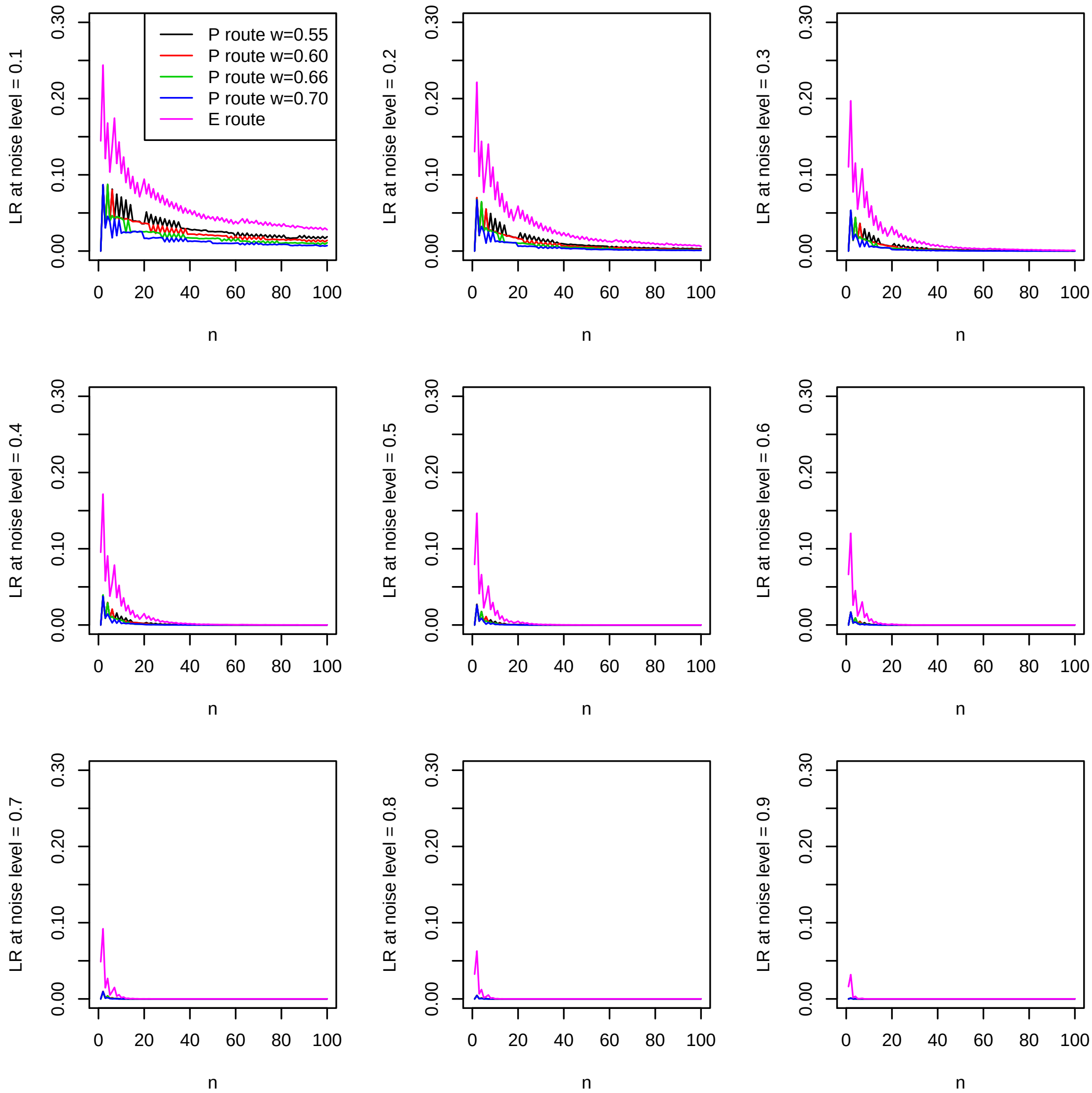

Figure 5.33: With average market condition, LR of both trading routes 2.3 and 2.4 under different noise levels p_2 for a put option expired data at $T = 100$.

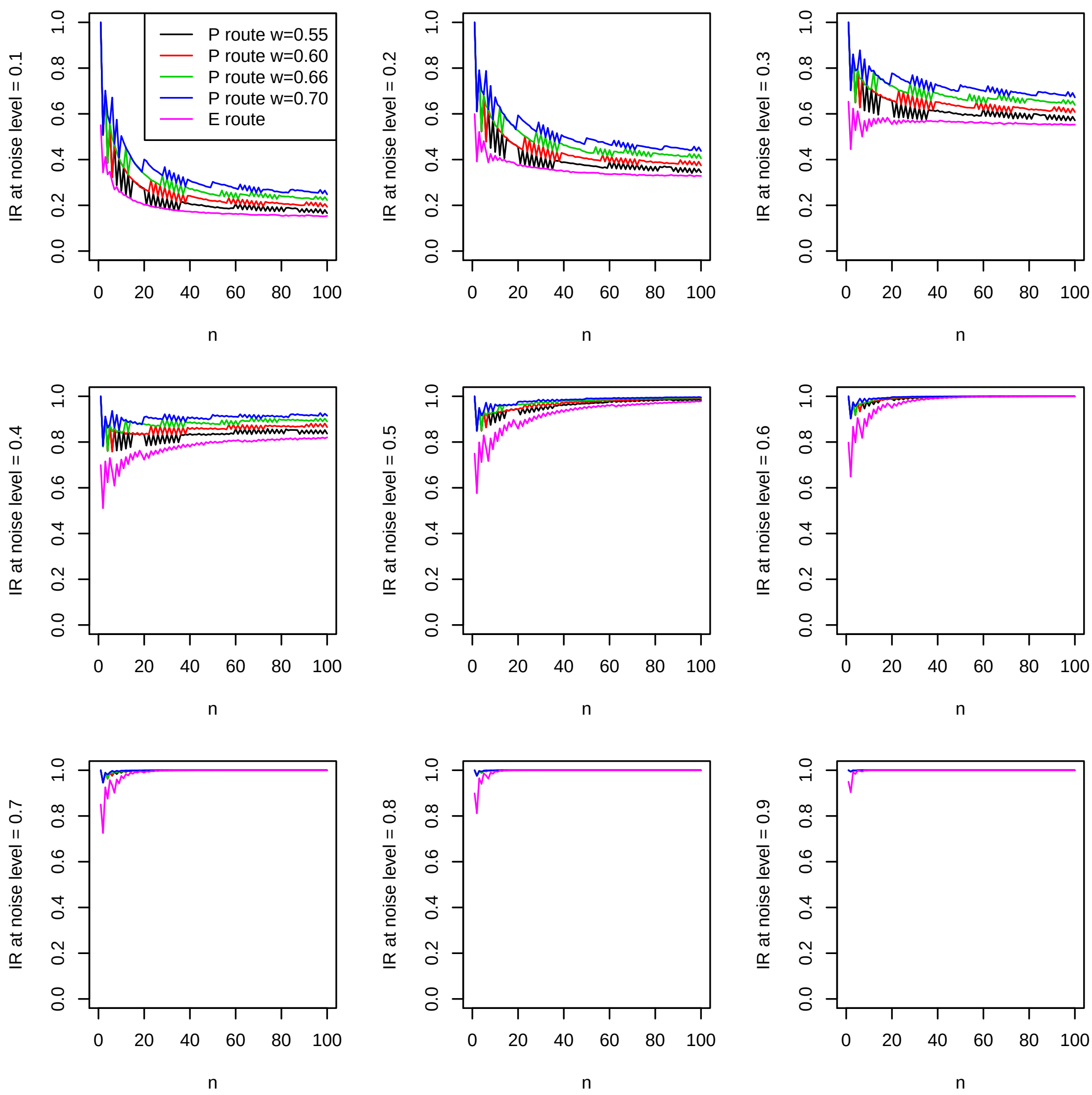

Figure 5.34: With average market condition, IR of both trading routes 2.3 and 2.4 under different noise levels p_2 for a put option expired data at $T = 100$.

As one may have expected, from Figure 5.34, the inaction rate for both 2.3 and 2.4 increase as the noise level increase. When noise level $p_2 \geq 0.5$, with sufficient data become available, the inaction rate of both trading routes increases to 1. When the data is only affected by low level noise $p_2 \in (0.1, 0.3)$, both trading routes are still able to learn useful information from the data. In these cases, both trading routes stay moderately active and take correct accordingly.(Indicated by the positive $\overline{f_i^P}$ surface in Figure 5.31). This again confirms the noise recognition capability of both trading routes.

Under the average market condition and throughout all different noise levels, both of trading routes perverse their original trading objectives respectively. In other words, route 2.3 has better risk control in R_{lr}^{P} while route 2.4 has greater $\overline{f_i^P}$ value which could be seen in Figure 5.33 and Figure 5.31 respectively.

Performance of NPI put option trading routes under different market conditions given imprecise data available

Given imprecise data, the performances of European put option trading routes are further evaluated under different market conditions below.

It is observed from simulations that both NPI European put option trading routes effectively and efficiently recognize the noise from the imprecise data and gradually take less trading action as the noise level increase. Moreover, when the noise level is low, both trading routes are able to extract the information of the underlying asset's market condition and execute correct action accordingly. Since both NPI put trading routes share similar decaying pattern in $\overline{f_i^P}$ and one complete example of $\overline{f_i^P}$ surface for a trading route requires nine pages of space, for the sake of brevity, we only present one complete example of average present value $\overline{f_i^P}$ surface for trading route 2.3 with threshold value $w = 0.6$ in the Appendix C. (See Figure C.1-C.9)

In order to have a direct comparison of route 2.3 and route 2.4, under market condition $p_1 = 0.1$, for a put option expiration data $T = 100$, the performance indexes of $\overline{f_i^A}$, R_{wr}^{A}, R_{lr}^{A} and R_{ir}^{A} for both put trading routes are plotted again each other in Figure 5.35-5.38.

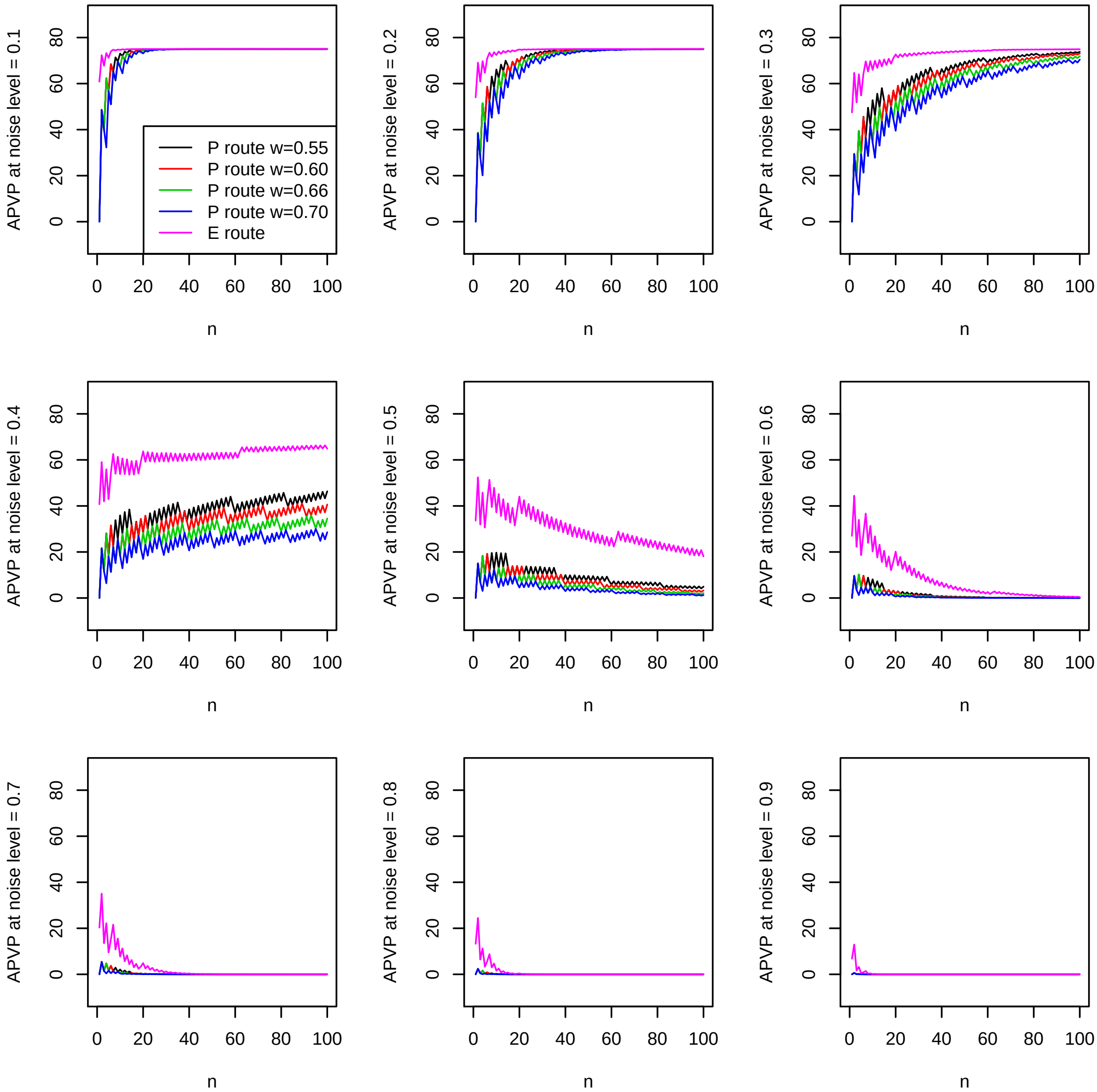

Figure 5.35: APVP of routes 2.3 and 2.4 at $T = 100$ under market condition $p_1 = 0.1$ and different noise levels p_2.

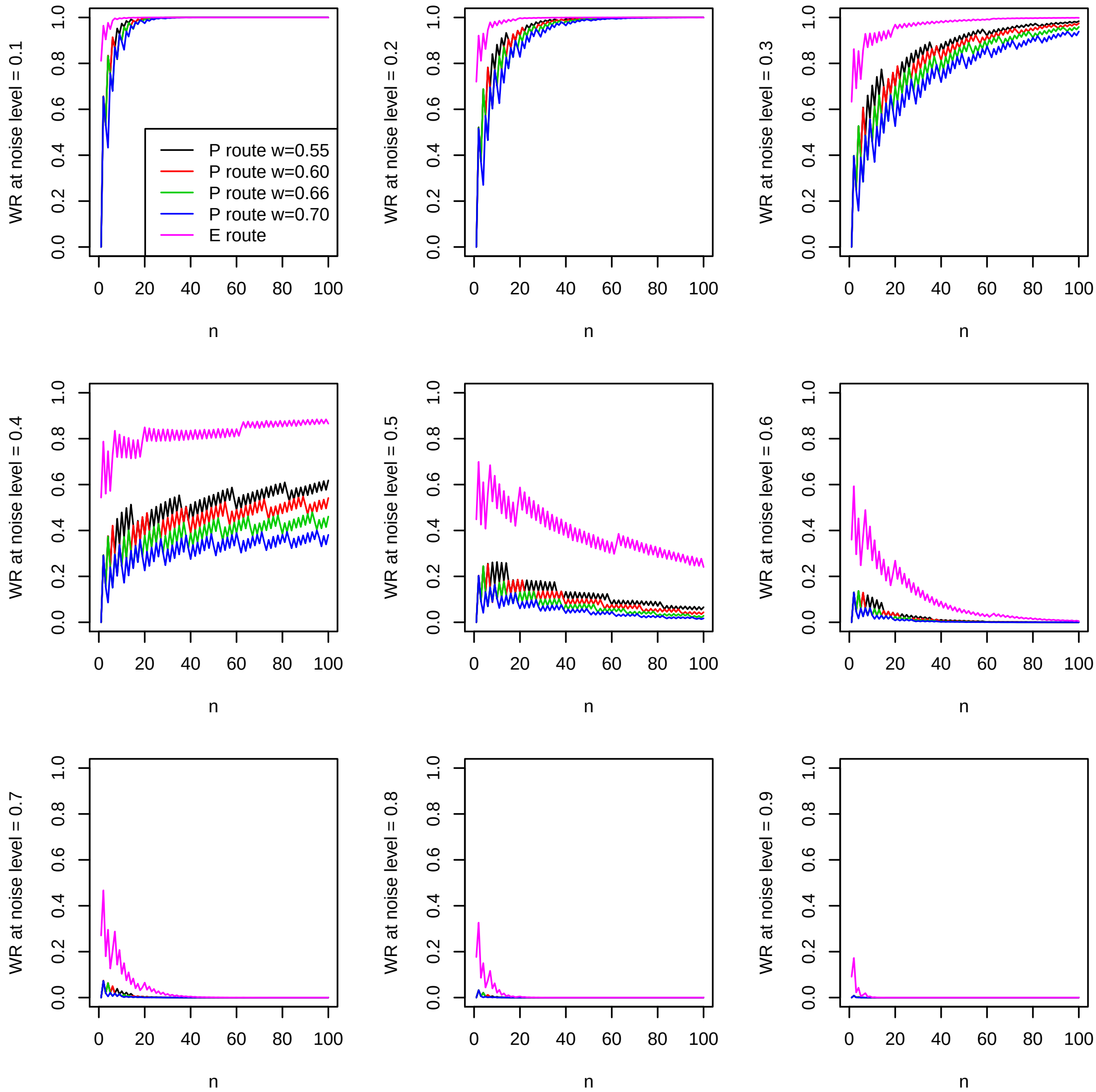

Figure 5.36: WR of routes 2.3 and 2.4 at $T = 100$ under market condition $p_1 = 0.1$ and different noise levels p_2.

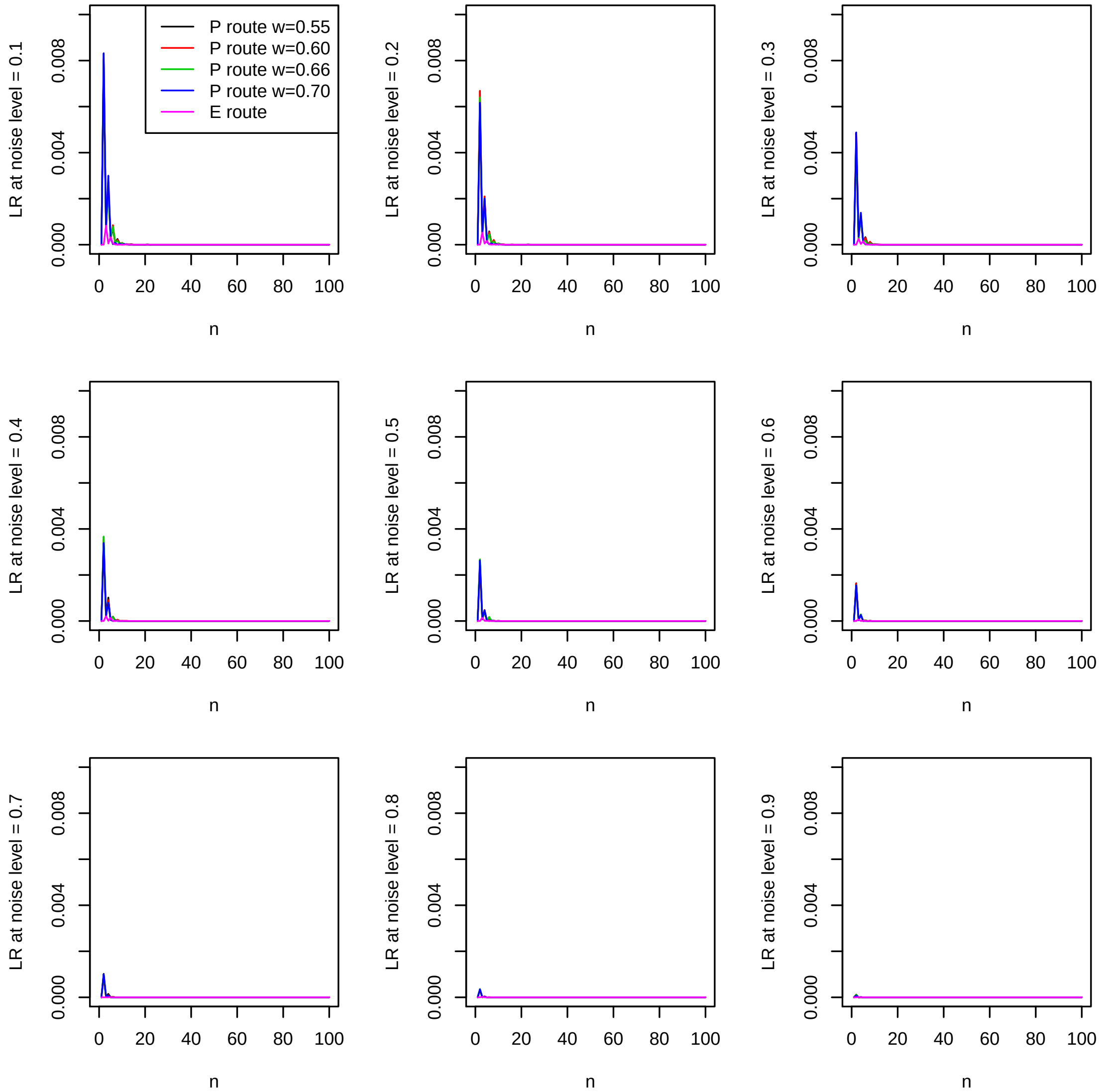

Figure 5.37: LR of routes 2.3 and 2.4 at $T = 100$ under market condition $p_1 = 0.1$ and different noise levels p_2.

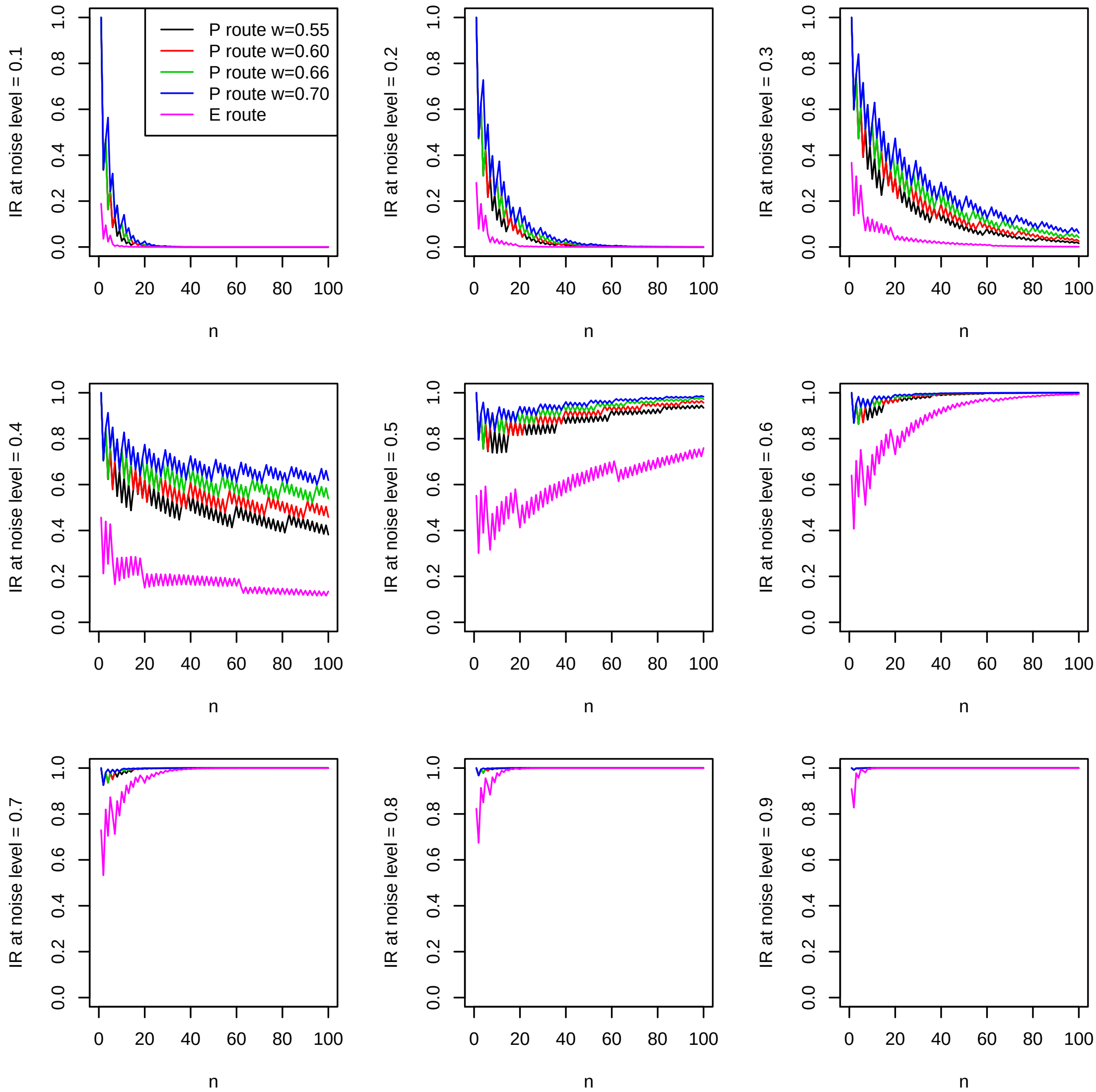

Figure 5.38: IR of routes 2.3 and 2.4 at $T = 100$ under market condition $p_1 = 0.1$ and different noise levels p_2.

From Figure 5.38, it could be found the inaction rates of both routes 2.3 and 2.4 increase, as the noise level increase in the data. More specifically, under the same noise level affection, given the same amount of data, route 2.3 has higher inaction rate than route 2.4. Although from Figure 5.37, it seems that route 2.3 and route 2.4 have similar loss rates under different noise levels affection, it is found in other simulation results that when market condition $p_1 \in (0.5, 0.9)$, under the same noise level affection, route 2.3 actually has a significantly lower loss rate than route 2.4. Therefore route 2.3 as imprecise probability trading route still has better risk control than route 2.4. From win rate profile in Figure 5.36, one may notice when noise level $p_2 \in (0.4, 0.6)$, the win rate of route 2.3 is significantly lower than route 2.4. However, as we have previously observed that both trading routes 2.3 and 2.4 have similar loss rate in noise region $p_2 \in (0.4, 0.6)$, the lower win rate of trading route 2.3 in noise region $p_2 \in (0.4, 0.6)$ is mainly attributed to its higher inaction rate as it tries to avoid making trading in the ambiguous situation. In contrast, from Figure 5.35, it could be found that route 2.4 generally has higher average present value payoff than route 2.3 under different noise levels.

To sum up, both proposed put option trading route 2.3 and route 2.4 are able to recognize noise level under different market conditions. They have their unique advantages in trading. While route 2.3 is more sensitive to noise and have good control of its loss rate, route 2.4 is better at producing higher present value payoff in the long run.

5.3 Overall review of European put option trading simulation

From Section 5.2.2, one could reach following conclusion:

The proposed NPI European put trading routes 2.3 and route 2.4 have decent performance under all market conditions and different noise levels.

Both trading routes are able to extract underlying information from the data effectively and efficiently and take correct action according to different market conditions preceived. The data learning process also has moderate noise resistance when only low level noise is presented. When the data is affected by high level noise, both trading routes are able to quickly recognize and stop taking any non sensible action.

Under no noise or low noise condition, given sufficient data presented, throughout all different market conditions, route 2.3 has better risk control while route 2.4 is able to achieve higher average present value payoff.

Chapter 6

Application of NPI method in portfolio assessment

In this chapter, consider several portfolios each of which comprises of assets and European options defined in Chapter 2, NPI for Bernoulli data is applied to do portfolio assessment with respect to two criteria.

6.1 Portfolio assessment scenario and criteria

Under binomial tree model, we consider portfolio assessment scenario as follow: one attempted to assess the profitability of multiple portfolios at time T. Each portfolio may contain assets or European options expired at time T. And all the components within each portfolio are independent. One tries to assess each portfolio by two different criteria as below:

Criterion I : Expected growth rate of each portfolio at time T

Criterion II: The probability which each portfolio generated a threshold amount of profit λ at time T

Denote the value of portfolio g at time T as a random variable Y_T^g. Y_T^g is then a weighted sum of its components value at time T. Denote the ith component's value of portfoilo g at time T as a random variable $X_T^{g,i}$, where $X_T^{g,i}$ can be asset $X_T^{g,i} = A_T^{g,i}$, or European call option $X_T^{g,i} = (A_T^{g,i} - K)^+$, or European put option $X_T^{g,i} = (K - A_T^{g,i})^+$. Also denote the weight of ith component in the portfolio g as

$w^{g,i}$ and there are in total N^g components in the portfolio g, then:

$$Y_T^g = \sum_{i=1}^{N^g} w^{g,i} X_T^{g,i}$$
$$\sum_{i=1}^{N^g} w^{g,i} = 1 \text{ and } 0 \leq w^{g,i} \leq 1$$

Therefore, the portfolio assessment criteria can be rewrite as follow:

Criterion I: Expected growth rate $E(r_T^g)$ of the portfolio g at time T

$$\begin{aligned} E(r_T^g) = \frac{\ln(\frac{E(Y_T^{g,i})}{Y_{0,i}^g})}{T} &= \frac{\ln(\frac{E(\sum_{i=1}^{N^g} w^{g,i} X_T^{g,i})}{\sum_{i=1}^{N^g} w^{g,i} x_T^{g,i}})}{T} \\ &= \frac{\ln(\frac{\sum_{i=1}^{N^g} w^{g,i} E(X_T^{g,i})}{\sum_{i=1}^{N^g} w^{g,i} x_T^{g,i}})}{T} \quad \text{By independence of components} \end{aligned}$$

Criterion II: The probability that the portfolio g generated a threshold amount of profit λ at time T. $p(Y_T^g > \lambda)$.

Let $e_T^{g,i}$ be any possible value that component i in the portfolio g could take at time T. then:

$$\begin{aligned} p(Y_T^g > \lambda) &= p(\sum_{i=1}^{N^g} w^{g,i} X_T^{g,i} > \lambda) \\ &= \sum_{e_T^{g,2}, e_T^{g,3}, \ldots e_T^{g,N^g}} p(w^{g,1} X_T^{g,1} > \lambda - \sum_{i=2}^{N^g} e_T^{g,i}) \prod_{i=2}^{i=N_g} p(w^{g,i} X_T^{g,i} = e_T^{g,i}) \end{aligned}$$

6.2 NPI method in portfolio assessment

Under above scenario setting, we apply NPI method to have a more conservative portfolio assessment by replacing the expectation operator and probability operator in criteria I & II to the lower expectation operator and lower probability operator respectively. Thus the criteria become :

Criterion I: Lower expected growth rate $\underline{E}(r_T^g)$ of portfolio g at time T

$$\underline{E}(r_T^g) = \frac{\ln(\frac{\underline{E}(Y_T^{g,i})}{Y_{0,i}^g})}{T} = \frac{\ln(\frac{\underline{E}(\sum_{i=1}^{N^g} w^{g,i} X_T^{g,i})}{\sum_{i=1}^{N^g} w^{g,i} x_T^{g,i}})}{T}$$

$$= \frac{\ln(\frac{\sum_{i=1}^{N^g} w^{g,i} \underline{E}(X_T^{g,i})}{\sum_{i=1}^{N^g} w^{g,i} x_T^{g,i}})}{T} \quad \text{By Formulas 2.1.10}$$

Criterion II: The lower probability that portfolio g generated a threshold amount of profit λ at time T. $p(Y_T^g > \lambda)$

Let $e_T^{g,i}$ be any possible value that component i in the portfolio g could take at time T. then:

$$\underline{p}(Y_T^g > \lambda) = \inf_{p \in P_m} p(\sum_{i=1}^{N^g} w^{g,i} X_T^{g,i} > \lambda) \quad \text{By Formulas 2.1.13}$$

$$= \sum_{e_T^{g,2}, e_T^{g,3}, \ldots e_T^{g,N^g}} p_{\underline{E}(X_T^{g,1})}(w^{g,1} X_T^{g,1} > \lambda - \sum_{i=2}^{N^g} e_T^{g,i}) \prod_{i=2}^{i=N_g} p_{\underline{E}(X_T^{g,i})}(w^{g,i} X_T^{g,i} = e_T^{g,i})$$

6.3 Simulation of NPI for portfolio assessment

In order to evaluate the performance of NPI method in portfolio assessment, computer simulation is conducted in this section using statistical software R version 3.5.1.

6.3.1 Simulation design

The simulation is designed as the following:

1. One generates a sequence of portfolios, each portfolio g contains a different combination of independent assets and European option with individual specified parameters. The number of component in each portfolio is set at 3 to 5 at random. Also, there are n historical data points for each component in each portfolio.

2. One then applies NPI for Bernoulli data to learn information from n historical data points from each component and induce 3 to 5 independent imprecise prob-

ability spaces for each portfolio. One subsequently find the $\underline{E}(r_T^g)$ and $\underline{p}(Y_T^g > \lambda)$ from NPI induced imprecise probability spaces for each portfolio g and produce a rank list for both criteria.

3. Evaluate the performance of NPI method for portfolio assessment of both criteria by comparing the true rank list and the rank list produced by NPI.

The comparison of rank lists is achieved by evaluating the average value of Error function $Err(R_v(E))$ and $Err(R_v(p))$. Error function $Err(R_v(E))$ and $Err(R_v(p))$ are defined below:

We denote true rank of portfolio g with respect to criterion I $E(r_T^g)$ and criterion II $p(Y_T^g > \lambda)$ as $R_v(E(r_T^g))$ and $R_v(p(Y_T^g > \lambda))$ in vth simulation trial. And we also denote the corresponding rank produced NPI method of portfolio g with respect to two criteria as $R_v(\underline{E}(r_T^g))$ and $R_v(\underline{p}(Y_T^g > \lambda))$ in vth simulation trial. We then define error function $Err(R_v(E))$ for the rank with respect to criterion I in vth trial as:

$$Err(R_v(E)) = \sum_g |R_v(E(r_T^g)) - R_v(\underline{E}(r_T^g))| \tag{6.3.1}$$

We also define error function $Err(R_v(p))$ for the rank with respect to criterion II in vth trial as:

$$Err(R_v(p)) = \sum_g |R_v(p(Y_T^g > \lambda)) - R_v(\underline{p}(Y_T^g > \lambda))| \tag{6.3.2}$$

Throughout the whole simulation section, we choose the λ in criterion II as $\lambda = e^{rT} Y_0^g$. Namely one is finding the lower probability of the event that portfolio g at time T would generate interest rate r^g that is greater than risk free rate r and the risk free interest rate r is set at $r = 0.003$.

6.3.2 Sample simulation trials

To have a better understanding of how the simulation conducted, we provide two simulation trials with risk free interest rate $r = 0.003$ for illustration. Denote $p^{g,i}$ as market condition for component i within portfolio g, and also denote $(n^{g,i}, j^{g,i})$ as Bernoulli data for component i within portfolio g.

Simulation trial 1: historical data point $n = 100$ in each component, assessment for future time $T = 10$

Table 6.1: Random generated portfolios in simulation trial 1

Portfolio g	$X^{g,i}_{T=10}$ within $Y^{g,i}_{T=10}$	$p^{g,i}$ of $X^{g,i}_{T=10}$	Data $(n^{g,i}, j^{g,i})$
1	$X^{1,1}_{T=10}$: A, 1.0487, 0.9438, 105	0.0378	(100,2)
	$X^{1,2}_{T=10}$: C, 1.0555, 0.9480, 109, 122	0.9544	(100,94)
	$X^{1,3}_{T=10}$: A, 1.0136, 0.9161, 97	0.1481	(100,14)
	$X^{1,4}_{T=10}$: A, 1.0067, 0.9921, 85	0.3249	(100,33)
	$X^{1,5}_{T=10}$: C, 1.0522, 0.9235, 97, 116	0.6147	(100,58)
2	$X^{2,1}_{T=10}$: P, 1.0354, 0.9577, 91, 81	0.5196	(100,57)
	$X^{2,2}_{T=10}$: P, 1.0883, 0.9539, 105, 92	0.5332	(100,52)
	$X^{2,3}_{T=10}$: A, 1.0042, 0.9165, 105	0.8379	(100,83)
3	$X^{3,1}_{T=10}$: P, 1.0317, 0.9148, 107, 90	0.2473	(100,21)
	$X^{3,2}_{T=10}$: A, 1.0461, 0.9379, 102,	0.0703	(100,5)
	$X^{3,3}_{T=10}$: P, 1.0326, 0.9863, 110 ,98	0.2695	(100,25)
4	$X^{4,1}_{T=10}$: C, 1.0781, 0.9698, 105, 115	0.9596	(100,97)
	$X^{4,2}_{T=10}$: A, 1.0383, 0.9155, 103	0.1068	(100,2)
	$X^{4,3}_{T=10}$: C, 1.0893, 0.9803, 107 ,116	0.0430	(100,7)
5	$X^{5,1}_{T=10}$: A, 1.0987, 0.9573, 93	0.2044	(100,25)
	$X^{5,2}_{T=10}$: P, 1.0818, 0.9681, 106 ,97	0.0863	(100,16)
	$X^{5,3}_{T=10}$: P, 1.0775, 0.9148, 101 ,80	0.2571	(100,29)

Table 6.2: NPI assessment for criteria I against the true value in simulation trial 1

Portfolio g	True $E(r^g_T)$	NPI $\underline{E}(r^g_T)$	rank of $E(r^g_T)$	rank of $\underline{E}(r^g_T)$
1	-0.0159	-0.0185	1	1
2	-0.0111	-0.0129	3	3
3	-0.0151	-0.0143	2	2
4	0.0266	0.0253	5	5
5	0.0133	0.0100	4	4

Table 6.3: NPI assessment for criteria II against the true value in simulation trial 1

Portfolio g	True $p(Y_T^g > \lambda)$	NPI $\underline{p}(Y_T^g > \lambda)$	rank of $p(Y_T^g > \lambda)$	rank of $\underline{p}(Y_T^g > \lambda)$
1	0.0039	0.0017	1	1
2	0.188	0.1653	3	3
3	0.0119	0.0113	2	2
4	0.7381	0.7390	5	5
5	0.7131	0.6112	4	4

Table 6.1 is a random portfolios generation table. In this simulation trial, it shows that one is given 5 portfolio g in column 1. In column 2, it gives the information of the components in each portfolio. Each individual portfolio g contains 3 to 5 component $X_{T=10}^{g,i}$ of which maybe asset, call or put. And the component $X_{T=10}^{g,i}$ information are presented in order of Type ("A"= Asset, "C" = call, "P"= Put), value of upward movement coefficient u, value of downward movement coefficient d, value of initial asset price $a_0^{g,i}$, value of strike price $K^{g,i}$ (if applicable). In column 3, the market condition $p^{g,i} \in (0,1)$ of each component $X_{T=10}^{g,i}$ within portfolio g is presented which the probability of upward movement in each time step, Also this information is hidden for one who tries to assess the portfolio. In column 4, one is given Data (n_i^g, j_i^g), namely the number of upward movement happened in past 100 historical time steps for component i in portfolio g

The simulation firstly calculate Y_0^g, the value of each portfolio at time 0. which is sum of each component $X_0^{g,i}$ at time 0. If $X^{g,i}$ is an asset, then $X_0^{g,i} = A_0^{g,i}$. If $X^{g,i}$ is an European option, one use the Q measure $q = \frac{(e^r - d)}{u-d}$ in CRR model to calculate non abitrage price as the the market price at time 0 and use it as the value of $X_0^{g,i}$. The simulation then calculate the true $E(r_T^g)$ and true $p(Y_T^g > \lambda)$ using all the information from table 6.1 and also calculate NPI $\underline{E}(r_T^g)$ and $\underline{p}(Y_T^g > \lambda)$ using the all the information from table 6.1 except the column 3. The simulation then produce the rank of $E(r^g)$ and $\underline{E}(r^g)$, also the rank $p(Y_T^g > \lambda)$ and $\underline{p}(Y_T^g > \lambda)$. This is presented in Table 6.2 and Table 6.3.

Finally, the simulations calculate the Error function value G and F for this simulation trial and store this value for later average value calculation.

From the simulation trial 1 in Tables 6.1-6.3, one could observe that both NPI lower operator $\underline{E}(r_T^g)$ and $\underline{p}(Y_T^g > \lambda)$ tend to produce a conservative lower value when one compare them with the true value of $E(r_T^g)$ and true $p(Y_T^g > \lambda)$. More importantly, it gives the correct ranking of the portfolios with respected to two criteria in this simulation trial and thus have zero value for both Error function $Err(R_v(E))$ and $Err(R_v(p))$.

Simulation trial 2: historical data point $n = 10$ in each component, assessment for future time $T = 10$

Table 6.4: Random generated portfolios in simulation trial 2

Portfolio g	$X_{T=10}^{g,i}$ within $Y_{T=10}^{g,i}$	$p^{g,i}$ of $X_{T=10}^{g,i}$	Data $(n^{g,i}, j^{g,i})$
1	$X_{T=10}^{1,1}$: A, 1.0793, 0.9565, 81	0.4655	(10,5)
	$X_{T=10}^{1,2}$: A, 1.0401, 0.9541, 106	0.4003	(10,4)
	$X_{T=10}^{1,3}$: C, 1.0460, 0.9740, 90, 93	0.5225	(10,8)
	$X_{T=10}^{1,4}$: C, 1.0973, 0.9095, 90, 101	0.7213	(10,6)
	$X_{T=10}^{1,5}$: P, 1.0981, 0.9815, 87, 74	0.9257	(10,9)
2	$X_{T=10}^{2,1}$: C, 1.0303, 0.9283, 91, 100	0.5679	(10,8)
	$X_{T=10}^{2,2}$: A, 1.0450, 0.9236, 105	0.0312	(10,0)
	$X_{T=10}^{2,3}$: P, 1.0661, 0.9430, 94, 79	0.7696	(10,7)
3	$X_{T=10}^{3,1}$: A, 1.0114, 0.9217, 103	0.4680	(10,5)
	$X_{T=10}^{3,2}$: C, 1.0992, 0.9363, 81, 95	0.3167	(10,4)
	$X_{T=10}^{3,2}$: A, 1.0789, 0.9854, 93	0.5077	(10,5)
	$X_{T=10}^{3,3}$: A, 1.0817, 0.9590, 88	0.7494	(10,9)
4	$X_{T=10}^{4,1}$: C, 1.0301, 0.9926, 84, 87	0.1107	(10,0)
	$X_{T=10}^{4,2}$: P, 1.0618, 0.9741, 107, 93	0.3059	(10,2)
	$X_{T=10}^{4,2}$: C, 1.0174, 0.9176, 85, 101	0.6011	(10,6)
	$X_{T=10}^{4,3}$: P, 1.0304, 0.9898, 87, 71	0.8144	(10,9)
5	$X_{T=10}^{5,1}$: P, 1.0697, 0.9470, 103, 90	0.6758	(10,7)
	$X_{T=10}^{5,2}$: P, 1.0136, 0.9846, 100, 82	0.0981	(10,1)
	$X_{T=10}^{5,3}$: C, 1.0199, 0.9851, 97, 113	0.2457	(10,0)

Table 6.5: NPI assessment for criteria I against the true value in simulation trial 2

Portfolio g	True $E(r_T^g)$	NPI $\underline{E}(r_T^g)$	rank of $E(r_T^g)$	rank of $\underline{E}(r_T^g)$
1	0.1942	0.1443	4	4
2	-0.5430	-0.5225	3	2
3	0.1992	0.2410	5	5
4	-0.6063	-0.0737	2	3
5	-1.5814	-1.2832	1	1

Table 6.6: NPI assessment for criteria II against the true value in simulation trial 2

Portfolio g	True $p(Y_T^g > \lambda)$	NPI $\underline{p}(Y_T^g > \lambda)$	rank of $p(Y_T^g > \lambda)$	rank of $\underline{p}(Y_T^g > \lambda)$
1	0.0214	0.0311	4	4
2	2.16×10^{-10}	0	1	1
3	0.9402	0.9257	5	5
4	1.09×10^{-8}	5.69×10^{-7}	2	3
5	0.0001	0	3	2

Tables 6.4-6.6 presents another simulation trial which has less historical data available $n = 10$ than the previous simulation trial. And in this simulation trial, the error function value for criterion I is $Err(R_v(E)) = |4-4| + |3-2| + |5-5| + |2-3| + |1-1| = 2$ and the error function value for criterion II is $Err(R_v(p)) = |4-4| + |1-1| + |5-5| + |2-3| + |3-2| = 2$

6.3.3 Simulation results

Simulation has been conducted for 5, 10, 15 portfolios assessment. Each case is simulated for 10000 times using the generation process same as the simulation trials provided.

The simulation results are presented as follow:

Simulation results for criterion I

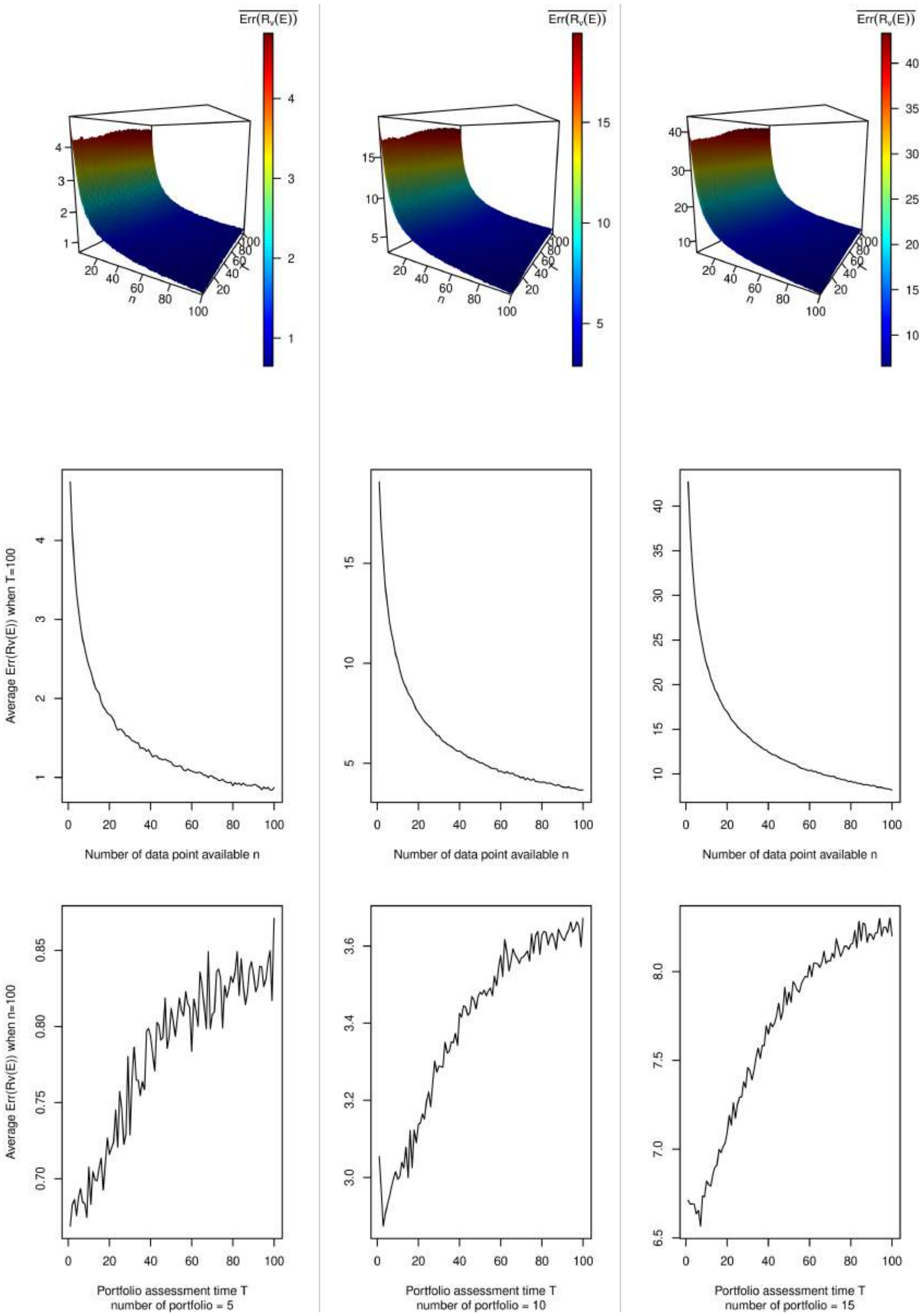

Figure 6.1: With 5, 10 ,15 portfolios, the average value of $Err(R_v(E))$ with different combinations of portfolio assessment time $T \in [1:100]$ and number of available data point$n \in [1, 100]$

Figure 6.1 presents the average value of criterion I error function $Err(R_v(E))$ with 5, 10, 15 portfolios assessment in each column. One should notice that in Figure 6.1, the z axis in the first row and the y axis in the second and third row are in different scales.

First row of Figure 6.1 presents $Err(R_v(E))$ with all combination of portfolio assessment time $T \in [1 : 100]$ and number of available data point $n \in [1, 100]$. From Figure 6.1 first row, one could see that as the number of portfolios increases, the average value of criterion I error function $Err(R_v(E))$ increase. However, since as the number of available data increase, the average value of criterion I error function $Err(R_v(E))$ decrease. It can be confirmed that the NPI portfolio assessment for criterion I effectively learn the information from the data. This could also be seen in the second row of Figure 6.1.

Second row of Figure 6.1 presents the average value of criterion I error function $Err(R_v(E))$ of 5, 10, 15 portfolios assessment in time $T = 100$ with different number of data point available $n \in [1, 100]$. As it could observed that for portfolios assessment in time $T = 100$, as number of data available increase, the average value of criterion I error function $Err(R_v(E))$ does decrease. The information learning speed is very quick when the number of data n is less than 37. When $n < 37$, on average, there are 0.0938, 0.3693, 0.8275 reduction in $Err(R_v(E))$ in the case of 5, 10, 15 portfolios assessment when one more unit data become available. Moreover, with sufficient data become available, the average value of $Err(R_v(E))$ for 5, 10, 15 portfolios assessment dropped to approximately 0.84, 3.66, 8.20 respectively. Those are acceptable value for average value of $Err(R_v(E))$ in the corresponding case. Because one should know that the smallest error could happens in the rank order is a permutation in any two adjacent true rank position and the smallest error will contribute a error value of 2 in $Err(R_v(E))$. Thus, the average values of $Err(R_v(E))$ of 0.84, 3.66, 8.20 in the case with 5, 10 ,15 portfolios are well performed results of NPI portfolio assessment for criterion I.

From third row of Figure 6.1, one can notice with the number of historical data $n = 100$ available, the average value of $Err(R_v(E))$ is insensitive to the portfolio assessment time T. As assessment time increase from 1 to 100, the increment of the average value of criterion I error function $Err(R_v(E))$ is less than 2 for all cases.

Simulation results for criterion II

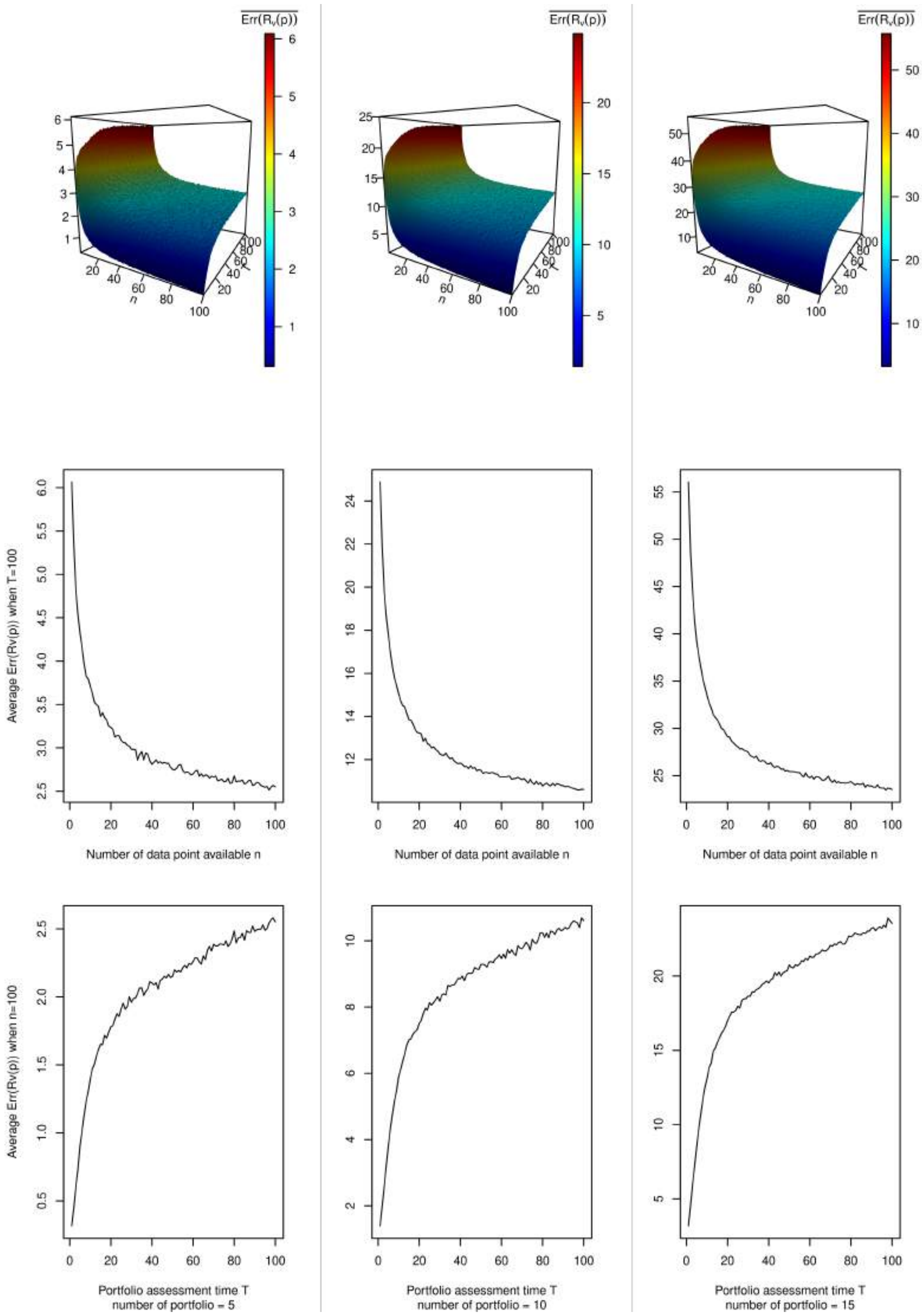

Figure 6.2: With 5, 10 ,15 portfolios, the average value of $Err(R_v(p))$ with different combinations of portfolio assessment time $T \in [1 : 100]$ and number of available data point$n \in [1, 100]$

Figure 6.2 presents the average value of criterion II error function $Err(R_v(p))$ with 5, 10, 15 portfolios assessment in each column. One should notice that in Figure 6.2, the z axis in the first row and the y axis in the second and third row are in different scales.

First row of Figure 6.2 presents $Err(R_v(p))$ with all combinations of portfolio assessment time $T \in [1 : 100]$ and the number of data available $n \in [1, 100]$. The surface of $Err(R_v(p))$ has different phenomenon than the case of $Err(R_v(E))$ in Figure 6.1. Although the proposed NPI assessment method with respect to criterion II generally has greater magnitude of error compared to NPI method with respect to criterion I, it is still able to learn information from the data and reduce the error as the number of available data point increase. Similar to $Err(R_v(E))$, the average magnitude of $Err(R_v(p))$ increase as the number of portfolios increase. It also appears that the proposed methods for criterion II is sensitive to the portfolio assessment time T when T is small. However, It gradually becomes insensitive when portfolio assessment time T become larger.

Second row of Figure 6.2 presents $Err(R_v(p))$ at portfolio assessment time $T = 100$ with different number of available data point $n \in [1, 100]$. When the number of available data point $n < 21$, the proposed method for criterion II has steep learning effect with 0.1426, 0.5843, 1.3495 decrement occurs in average value of $Err(R_v(p))$ in the case of 5, 10, 15 portfolios assessment when one more data point becomes available. After sufficient data is presented, the error function $Err(R_v(p))$ drop to roughly 2.51, 10.59, 23.48 for 5, 10, 15 portfolios assessment respectively. Although the performance of the proposed method for criterion II is less remarkable than the method for criterion I when portfolio assessment time T is large, it nevertheless provide a good solution in the cases when portfolio assessment time T is small, indicated by the third row of Figure 6.2.

Third row of Figure 6.2 presents average value of $Err(R_v(p))$ with different portfolio assessment time $T \in [1, 100]$ when the number of available data is $n = 100$. It could be seen that $Err(R_v(p))$ is very sensitive to portfolio assessment time when $T \leq 22$. When $T = 20$, given $n = 100$ unit of available data, $Err(R_v(p))$ has average value approximately at 1.78, 7.50, 17.01 for 5, 10, 15 portfolios assessments respectively. This indicated the proposed method for criterion II is still able to produce acceptable evaluation when portfolio assessment time T is small.

6.4 Conclusion of NPI in portfolio assessment

In this chapter, under the binomial tree model, NPI method is applied to portfolio assessment with respect to two criteria. Criterion I is the expected growth rate of a portfolio at a specific future time and Criterion II is the probability that portfolio generated a threshold amount of profit at a specific future time. Both criteria are mathematically formulated. Simulations are conducted to evaluate the performance of proposed NPI portfolio assessment with respect to the criteria. It confirmed the proposed NPI assessment methods are able to learn the information from the data and produce a good ranking list for a given set of portfolios. It should, however, be noticed that the proposed NPI method for criterion I work well for different assessment time while the proposed NPI method for criterion II only well performed when the assessment time is in the short future.

Chapter 7

Conclusions and Future Directions

This Chapter presents a summary of the results in this thesis. After that, some potential future research directions are suggested.

7.1 Conclusions

This thesis further developed imprecise probability methodology NPI for Bernoulli data to address two current challenging issues and the developed NPI for Bernoulli data is applied in finance with performance evaluations.

In Chapter 2, a set of imprecise probability definitions based on the concept of the mass function from Weichselberger's axiomatization of imprecise probability theory [39] and Dempster-Shafer's notion of basic probability assignment [26, 34] is introduced which provided the basic framework for further development of NPI for Bernoulli data in Chapter 3. After that, the imprecise probability methodology — Nonparametric predictive inference (NPI) is presented. Its current development of NPI for Bernoulli data is reviewed in detail in which two of its challenging issues are identified. Finally, to facilitate later NPI financial application in Chapter 4-6, relevant financial objects are defined and corresponding financial concepts and terminologies are explained.

In Chapter 3, NPI for Bernoulli data is further developed to address two major challenging issues—computation of imprecise expectation for a general function of multiple future stages observations and handling of imprecise data. To address the former, a GMA algorithm to find imprecise expectation measure for a general function of a finite random variable on an imprecise probability space is presented. With NPI latent variable representation, the mass function of NPI for Bernoulli data is constructed which enable the usage of GMA algorithm in NPI for Bernoulli data. A numerical example of how to use the GMA algorithm in NPI for Bernoulli data is provided. To address the latter, NPI's path counting method in the underlying lattice representation is extended which leads to the development of NPI for imprecise Bernoulli data. The property of NPI for imprecise Bernoulli data is identified and presented with a numerical example.

Chapter 4–6 are the sequels of Chapter 3 and Chapter 2. In Chapter 4, NPI for Bernoulli data and imprecise Bernoulli data are applied to asset trading under the binomial tree model in a presetting scenario. Two NPI based asset trading routes, one based imprecise probability, one based on imprecise expectation are proposed. Simulations are conducted to evaluate the performance of proposed asset trading routes under the different market conditions and noise levels in the data. From the simulation results, it is found that the proposed NPI asset trading routes are able to learn information from data effectively and have predictive property. Both of asset trading routes are also able to recognize noise contained in the data and adjust its strategy correspondingly based on the information it could learn from the noisy data. Under the average market condition, both asset trading routes are able to produce average positive present value payoff in the long run regardless of what noise level is contained in the data. It is also found that the proposed asset trading routes have different primary trading objectives. While the NPI imprecise expectation asset trading routes have greater average present value payoff generally, the NPI imprecise probability asset trading route has better risk control in the loss rate. Overall, the proposed NPI asset trading routes have good performance under the different market condition and noise level. Depending on one's risk preference in trading, one can choose or combine them to use.

In Chapter 5, under binomial tree model and using CRR non arbitrage price for European options as current market price, NPI for Bernoulli data and imprecise Bernoulli data are applied to European options trading in the prescribed scenarios.

Two NPI based European call option trading routes and two NPI based European put option trading routes are proposed. Simulations are conducted correspondingly to evaluate their performance under the different market condition and noise level in the data. From the simulation results, one could confirm that the proposed trading routes share similar trading primary objectives as appeared in Chapter 4. While imprecise expectation trading routes are able to produce better average present value payoff in the long run, imprecise probability trading routes are better at risk control in loss rate. It is also confirmed the proposed NPI trading routes for both European call option and put option effectively and efficiently learn the information from the data and are capable of executing correct action accordingly. When noise is presented in data, the proposed trading routes are able to recognize the noise level in the data. This is indicated by the gradual increment of inaction rate as noise level increases. Overall all proposed NPI European option trading routes are well performed under the different market condition and noise level in the data.

In Chapter 6, under the binomial tree model and with CRR non arbitrage European option price as current market price, NPI for Bernoulli data is applied to portfolio assessment with respect to two criteria. The criteria are firstly mathematically formulated. Subsequently, NPI method is applied. The performance of NPI portfolio assessment is evaluated via simulation. From the simulation results, it is confirmed that the NPI method of portfolio assessment for Criterion I (Expected grow rate of each portfolio at time T) could effectively learn information from the data. The average rank error of NPI method for Criterion I could reduce to a satisfactory level as the number of data increase. Also, the average rank error of NPI method for Criterion I is insensitive to the portfolio assessment time T. In contrast, although the NPI method of portfolio assessment for Criterion II (The probability which each portfolio generated a threshold amount of profit λ at time T) is also able to learn the information from the data effectively. However, the average rank error is in higher magnitude than the NPI method for Criterion I. Moreover, the NPI method for Criterion II is sensitive to assessment time T, when assessment time T become larger, NPI method for Criterion II tends to have a higher average rank error. Nevertheless, if assessment time T is small, NPI method for Criterion II is still able to produce a satisfactory rank for the portfolios given.

7.2 Future directions

The presented results from the thesis lead to several potential future research directions.

First, the lattice counting approach in the mass function construction of NPI for Bernoulli data could be adopted in the construction of the mass function in NPI for other data type. This would enable the computation for the imprecise expectations of a general function of multiple future stages observations in NPI for other data types.

Second, given a imprecise probability space $[\Omega, \mathscr{A}, m(\cdot)]$, a discrete random variable is a function $X : \Omega \longrightarrow F$. one can define the lower variance $\underline{V}$ and the upper variance $\overline{V}$ of X as:

$$\underline{V}(X) = \inf_{p(\cdot)\in P_m} \sum_{\omega\in\Omega} \left((X(w) - \sum_{\omega\in\Omega} X(\omega)p(\omega))^2 \times p(\omega) \right)$$
$$\overline{V}(X) = \sup_{p(\cdot)\in P_m} \sum_{\omega\in\Omega} \left((X(w) - \sum_{\omega\in\Omega} X(\omega)p(\omega))^2 \times p(\omega) \right)$$

One can also further define the lower variance measure $p_{\underline{V}(X)}(\cdot)$ and upper variance measure $p_{\overline{V}(X)}(\cdot)$ as

$$p_{\underline{V}(X)}(\cdot) = \operatorname*{arginf}_{p(\cdot)\in P_m} \sum_{\omega\in\Omega} \left((X(w) - \sum_{\omega\in\Omega} X(\omega)p(\omega))^2 \times p(\omega) \right)$$
$$p_{\overline{V}(X)}(\cdot) = \operatorname*{argsup}_{p(\cdot)\in P_m} \sum_{\omega\in\Omega} \left((X(w) - \sum_{\omega\in\Omega} X(\omega)p(\omega))^2 \times p(\omega) \right)$$

Finding the imprecise variance $p_{\underline{V}(X)}(\cdot)$ and $p_{\overline{V}(X)}(\cdot)$ in NPI for Bernoulli data is then an interesting research problem which involving solving a quadratic linear programming problem with the mass function developed in this thesis.

Third, Chapter 3 presents a way to compute imprecise expectation for a general function of multiple future stages observations S_T. One now may consider the trading scenario where one needs to trade a bundle of different European options with the same underlying asset and expiration date. The payoff of the bundle at future time T will then be a non-monotonic function of S_T and evaluating the performance of NPI trading routes in this new scenario could be an interesting topic.

Lastly, Chapter 6 in this thesis used NPI method to assess portfolio in terms of ranking by two criteria individually. There is also an opportunity for one used NPI method to do portfolio optimization with two criteria taken into accounting simultaneously.

Bibliography

[1] Aboalkhair, M.A. (2012). Nonparametric predictive inference for system reliability. *A thesis presented for the degree of Doctor of Philiosophy*, University of Durham.

[2] Arts, G.R.J., Coolen, F.P.A. and van der Laan, P. (2004). Nonparametric predictive inference in statistical process control *Quality Technology and Quantitative Management*, **1**, 201–216.

[3] Arts, G.R.J. and Coolen, F.P.A. (2008). Two nonparametric predictive control charts *Journal of Statistical Theory and Practice*, **2**, 499–512.

[4] Augustin, T. and Coolen, F.P.A. (2004). Nonparametric predictive inference and interval probability. *Journal of Statistical Planning and Inference*, **124**, 251–272.

[5] Baker, R.M., Coolen-Maturi, T. and Coolen, F.P.A. (2017) Nonparametric predictive inference for stock returns, *Journal of Applied Statistics*, **44**, 1333–1349.

[6] Boole, G. (1854). *An investigation of the laws of thought on which are founded the mathematical theories of logic and probabilities*. Walton and Maberley, London.

[7] Chen, J., Coolen, F.P.A. and Coolen-Maturi, T. (2019) On nonparametric predictive inference for asset and European option trading in the binomial tree model. *Journal of The Operational Research Society*, Under final review.

[8] Coolen, F.P.A (1998). Low structure imprecise predictive inference for Bayes' problem. *Statistics and Probability Letters*, **36**, 349–357.

[9] Coolen, F.P.A. and Yan, K.J. (2004) Nonparametric predictive inference with right-censored data. *Journal of Statistical Planning and Inference*, **126**, 25–54.

[10] Coolen-Schrijner, P. and Coolen, F.P.A. (2004). Adaptive age replacement based on nonparametric predictive inference. *Journal of the Operational Research Society* **55**, 1281–1297.

[11] Coolen, F.P.A. and Augustin, T. (2005) Learning from multinomial data: a nonparametric predictive alternative to the Imprecise Dirichlet Model. *Proceedings of the Fourth International Symposium on Imprecise Probability: Theories and Applications*, 125–134.

[12] Coolen, F.P.A. and Coolen-Schrijner, P. (2005). Nonparametric predictive reliability demonstration for failure-free periods. *IMA Journal of Management Mathematics*, **16**, 1–11.

[13] Coolen, F.P.A. (2006). On nonparametric predictive inference and objective Bayesianism. *Journal of Logic, Language and Information*, **15**, 21–47.

[14] Coolen, F.P.A. (2006). On probabilistic safety assessment in case of zero failures. *Journal of Risk and Reliability*, **220**, 105–114.

[15] Coolen, F.P.A. (2007). Nonparametric prediction of unobserved failure modes. *Journal of Risk and Reliability*, **221**, 207–216.

[16] Coolen, F.P.A. (2008). Nonparametric Predictive Inference for Bernoulli Quantities with Set-Valued Data. *In: Soft Methods for Handling Variability and Imprecision. Advances in Soft Computing*, **48**. Springer, Berlin, Heidelberg. 85–91.

[17] Coolen, F.P.A. and Augustin, T. (2009). A nonparametric predictive alternative to the Imprecise Dirichlet Model: the case of a known number of categories. *International Journal of Approximate Reasoning*, **50**, 217–230.

[18] Coolen, F.P.A. (2011). Nonparametric predictive inference. *In: International Encyclopedia of Statistical Science*, ed. M. Lovric, 968–970. Springer, Berlin.

[19] Coolen, F.P.A. and Al-nefaiee, A.H. (2012). Nonparametric predictive inference for failure times of systems with exchangeable components. Proceedings of the Institution of Mechanical Engineers, *Part O: Journal of Risk and Reliability.* **226(3)**: 262–273.

[20] Coolen-Maturi, T., Coolen-Schrijner, P. and Coolen, F.P.A. (2011). Nonparametric predictive selection with early experiment termination. *Journal of Statistical Planning and Inference* **141(4)**: 1403–1421.

[21] Coolen-Maturi, T., Coolen-Schrijner, P. and Coolen, F.P.A. (2011). Nonparametric predictive multiple comparisons of lifetime data. *Communications in Statistics - Theory and Methods* **41(22)**: 4164–4181.

[22] Coolen-Schrijner, P., Coolen, F.P.A. and Shaw, S.C. (2006). Nonparametric adaptive opportunity-based age replacement strategies. *Journal of the Operational Research Society* **57**, 63–81.

[23] Coolen-Schrijner, P. and Coolen, F.P.A. (2007). Nonparametric predictive comparison of success-failure data in reliability. *Journal of Risk and Reliability*, **221** , 319–327.

[24] Coolen-Schrijner, P., Maturi, T.A. and Coolen, F.P.A. (2009). Nonparametric predictive precedence testing for two groups, *Journal of Statistical Theory and Practice.* **3**: 273–287.

[25] Cox, J.C., Ross, S.A. and Rubinstein, M. (1979) Option pricing: A simplified approach. *Journal of Financial Economics*, **7**, 229–235.

[26] Dempster, A.P. (1967). Upper and lower probabilities induced by a multivalued mapping. *The Annals of Mathematical Statistics.* **38 (2)**, 325–339.

[27] He, T., Coolen, F.P.A. and Coolen-Maturi, T. (2019) Nonparametric predictive inference for European option pricing based on binomial tree model. *Journal of The Operational Research Society*

[28] He, T. (2019). Nonparametric predictive inference for option pricing based on the binomial tree model. *A thesis presented for the degree of Doctor of Philiosophy*, Univeristy of Durham.

[29] Hill, B.M. (1968). Posterior distribution of percentiles: Bayes' theorem for sampling from a population. *Journal of the American Statistical Association*, **63**, 677–691.

[30] Hill, B.M. (1988). De Finetti's Theorem, Induction, and $A_{(n)}$ or Bayesian nonparametric predictive inference(with discussion). *Bayesian Statistics 3*, Bernardo, *et al.* (Eds.). Oxford University Press, 211–241.

[31] Hill, B.M. (1993). Parametric models for $A_{(n)}$: splitting processes and mixtures. *Journal of the Royal Statistical Society B*, **55**, 423–433.

[32] Hull, J.C. (2009). *Options, Futures and other Derivatives* (7th ed.) Prentice-Hall.

[33] Keith, C. and Dirk, N. (2001). *Financial Engineering Derivatives and Risk Management* (7th ed.) John Wiley and Sons Ltd.

[34] Shafer, G. (1976). *A Mathematical Theory of Evidence.* Princeton University Press.

[35] Vicig, P. (2014). Financial risk measurement. *In: Introduction to Imprecise Probabilities, Chapter 12* ed. Augustin, T., Coolen, F.P.A., Cooman, G. and Troffaes, M.C.M., John Wiley and Sons Ltd.

[36] Walley, P. (1991). Statistical Reasoning with Imprecise Probabilities. *Chapman and Hall;* London (UK).

[37] Walley, P. (1996). Inferences from multinomial data:learning about a bag of marbles. *Journal of the Royal Statistical Society, Series B*, **58(1)**: 3–34.

[38] Walley, P. (1999). Towards a unified theory of imprecise probability. *International Journal of Approximate Reasoning*, **24** (2000), 125–148.

[39] Weichselberger, K. (2000). The theory of interval-probability as a unifying concept for uncertainty. *International Journal of Approximate Reasoning*, **24**, 149–170.

[40] Weichselberger, K. (2001). Elementare Grundbegriffe einer allgemeineren Wahrscheinlichkeitsrechnung I. Intervallwahrscheinlichkeit als umfassendes Konzept. *Physika; Heidelberg (Germany).*

[41] Weichselberger, K. (2005). The logical concept of probability and statistical inference. In Fabio G. Cozman, Robert Nau, and Teddy Seidenfeld, editors, *ISIPTA' 05: Proc. 4th Int. Symp. on Imprecise Probabilities and Their Applications.* 396–405.

Appendix A

Performance evaluation of NPI asset trading routes

This appendix presents a full example of average present value payoff surface $\overline{f_i^A}$ of asset trading route 1.1 with threshold parameter value $w = 0.6$ under the different market conditions and subject to the different noise levels. The example is part of simulation results in Chapter 4 which shows that the proposed NPI asset trading routes' noise recognition capability under different market conditions and data learning ability under low level noise affection. As mentions in Chapter 4, trading route 1.1 and 1.2 have similar decaying phenomenons in surface $\overline{f_i^A}$. The presented example could be regarded as the general property demonstrations for both route 1.1 and route 1.2.

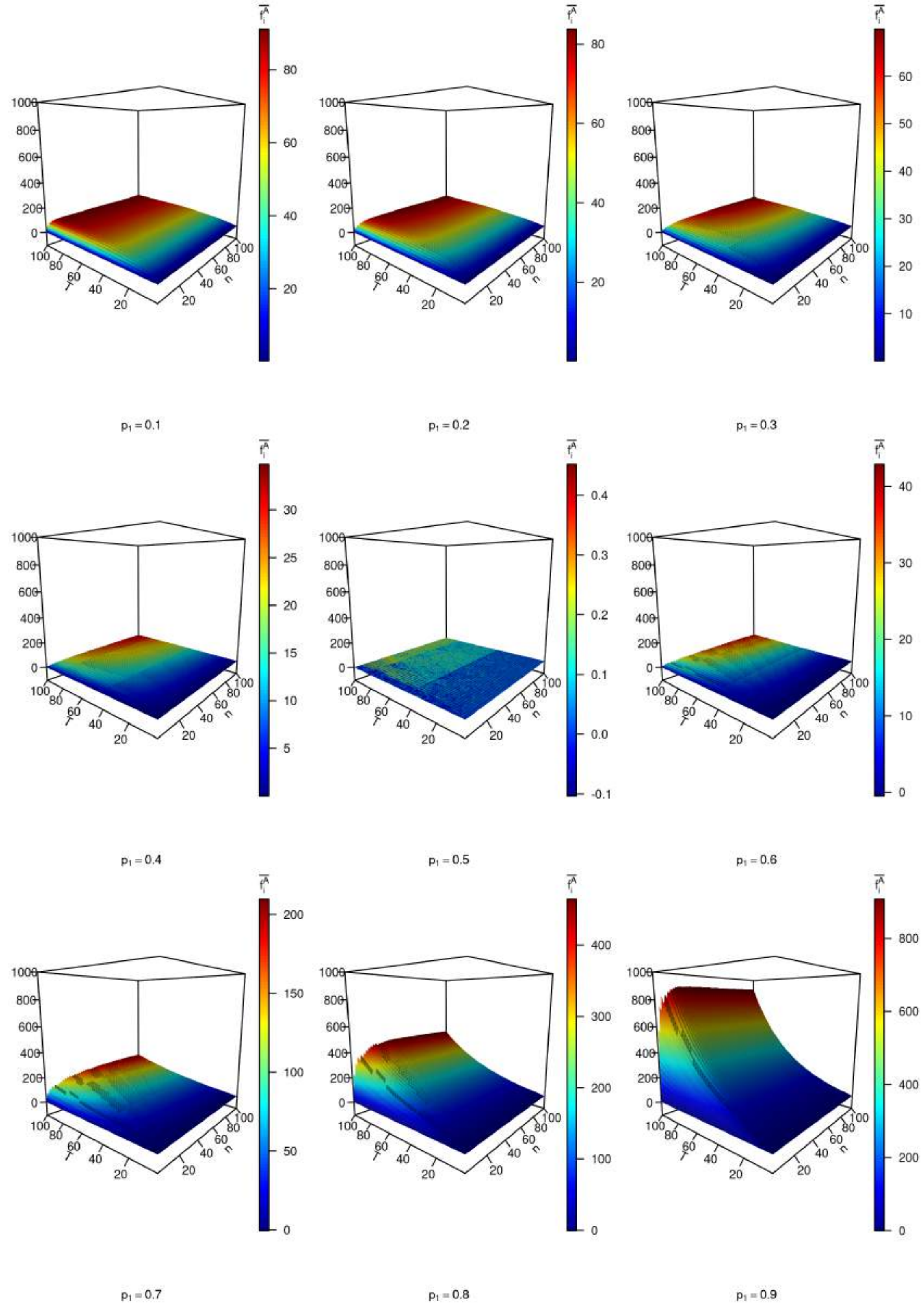

Figure A.1: Trading route 1.1, $w = 0.6$, noise $p_2 = 0.1$.

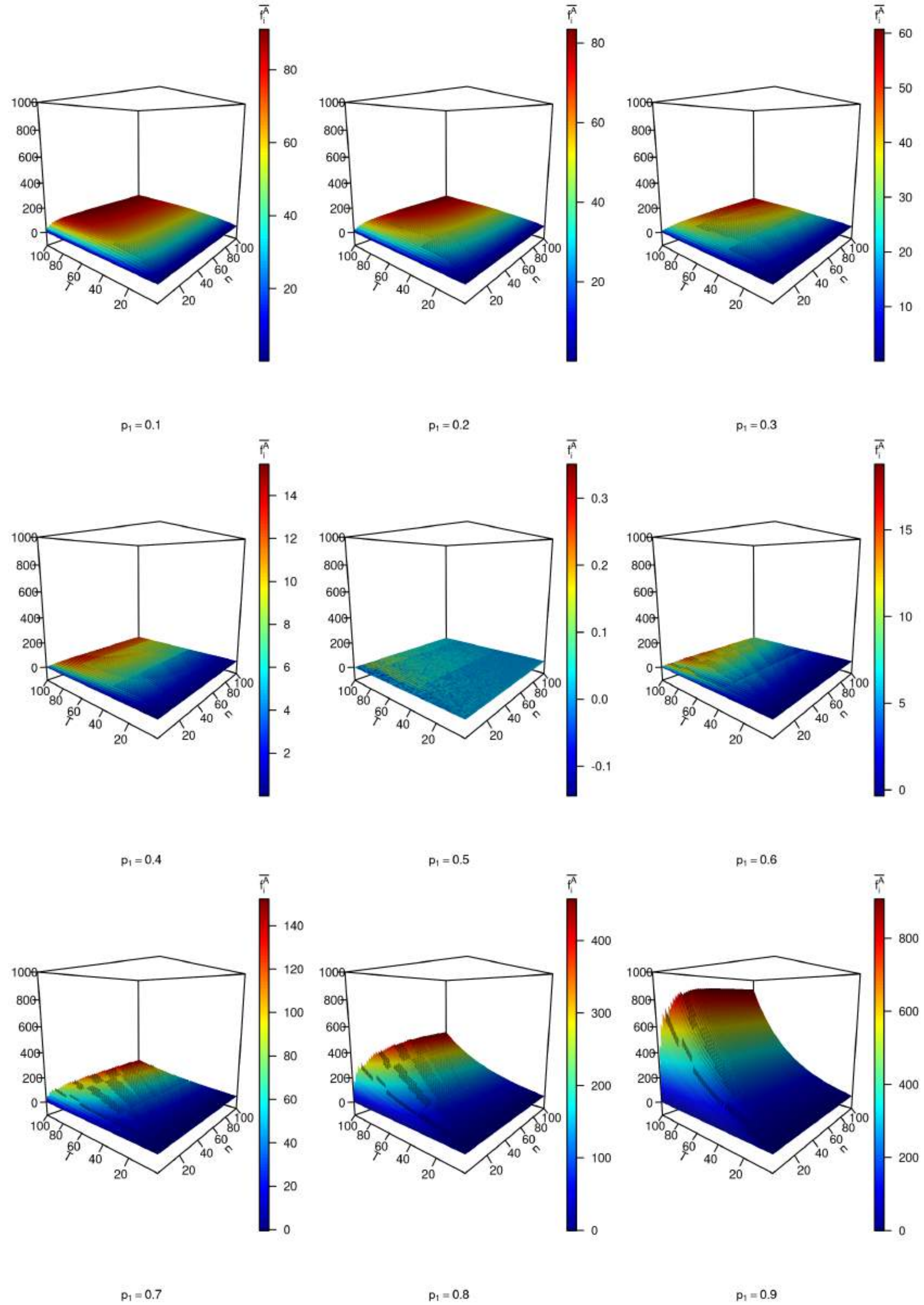

Figure A.2: Trading route 1.1, $w = 0.6$, noise $p_2 = 0.2$.

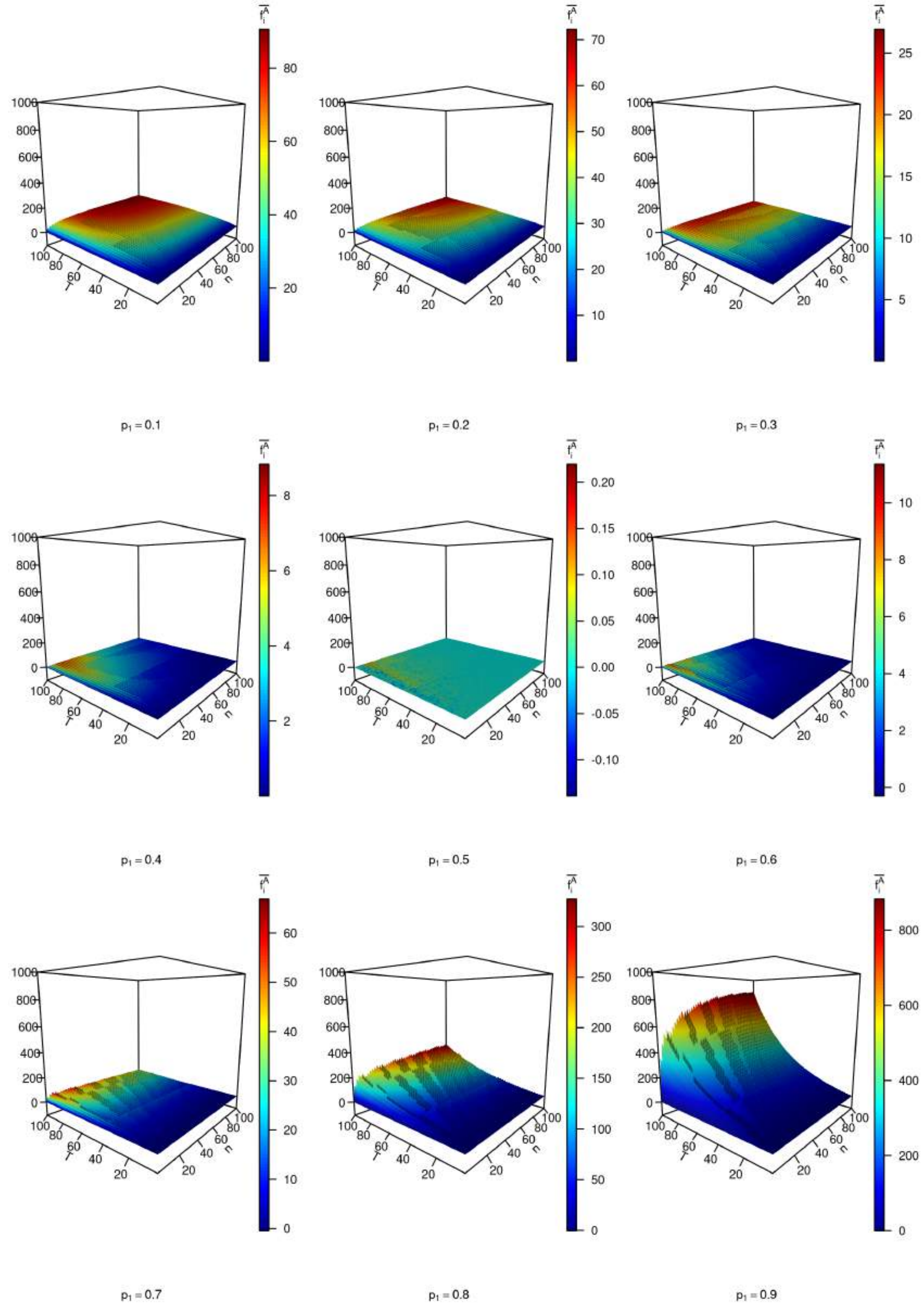

Figure A.3: Trading route 1.1, $w = 0.6$, noise $p_2 = 0.3$.

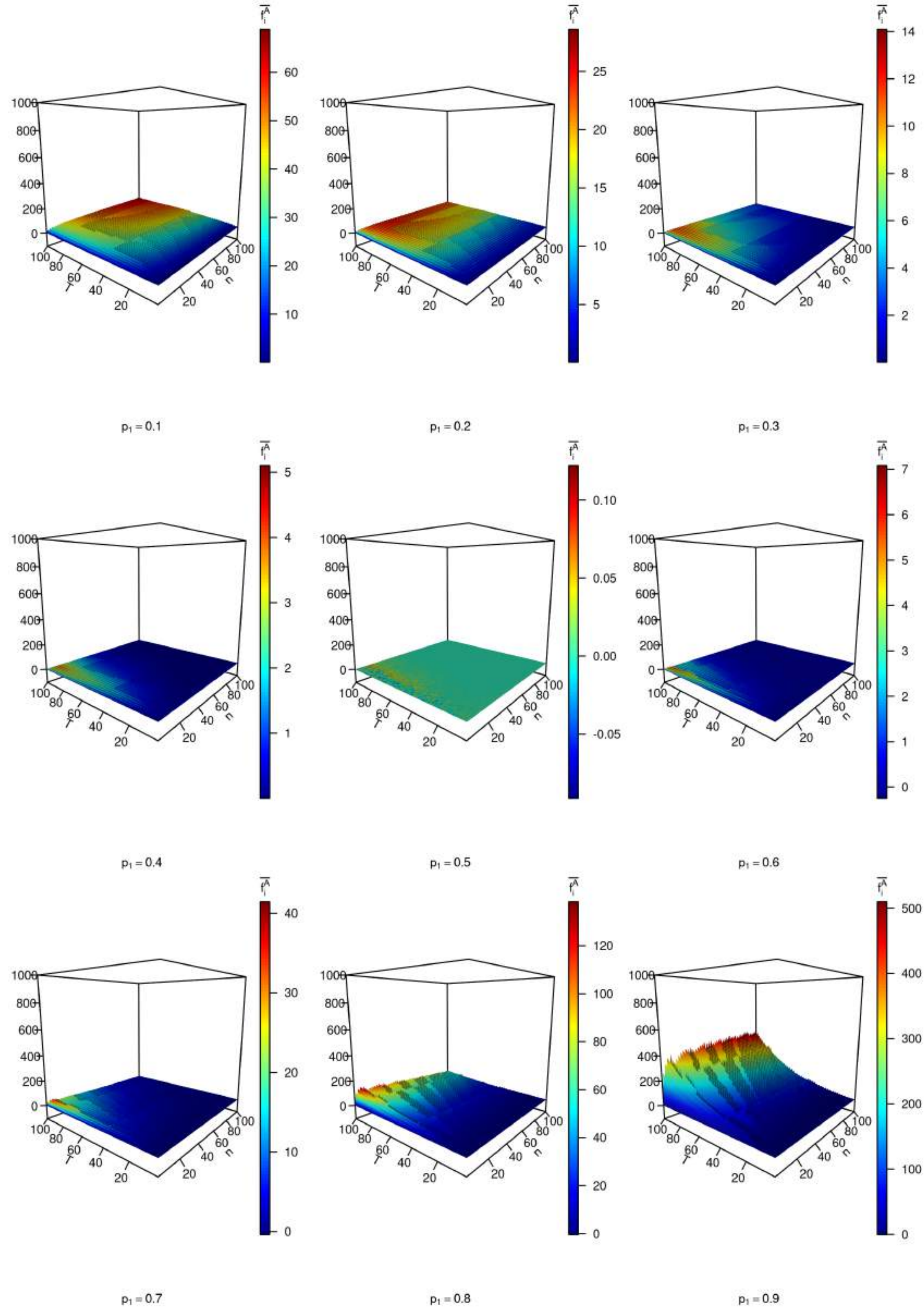

Figure A.4: Trading route 1.1, $w = 0.6$, noise $p_2 = 0.4$.

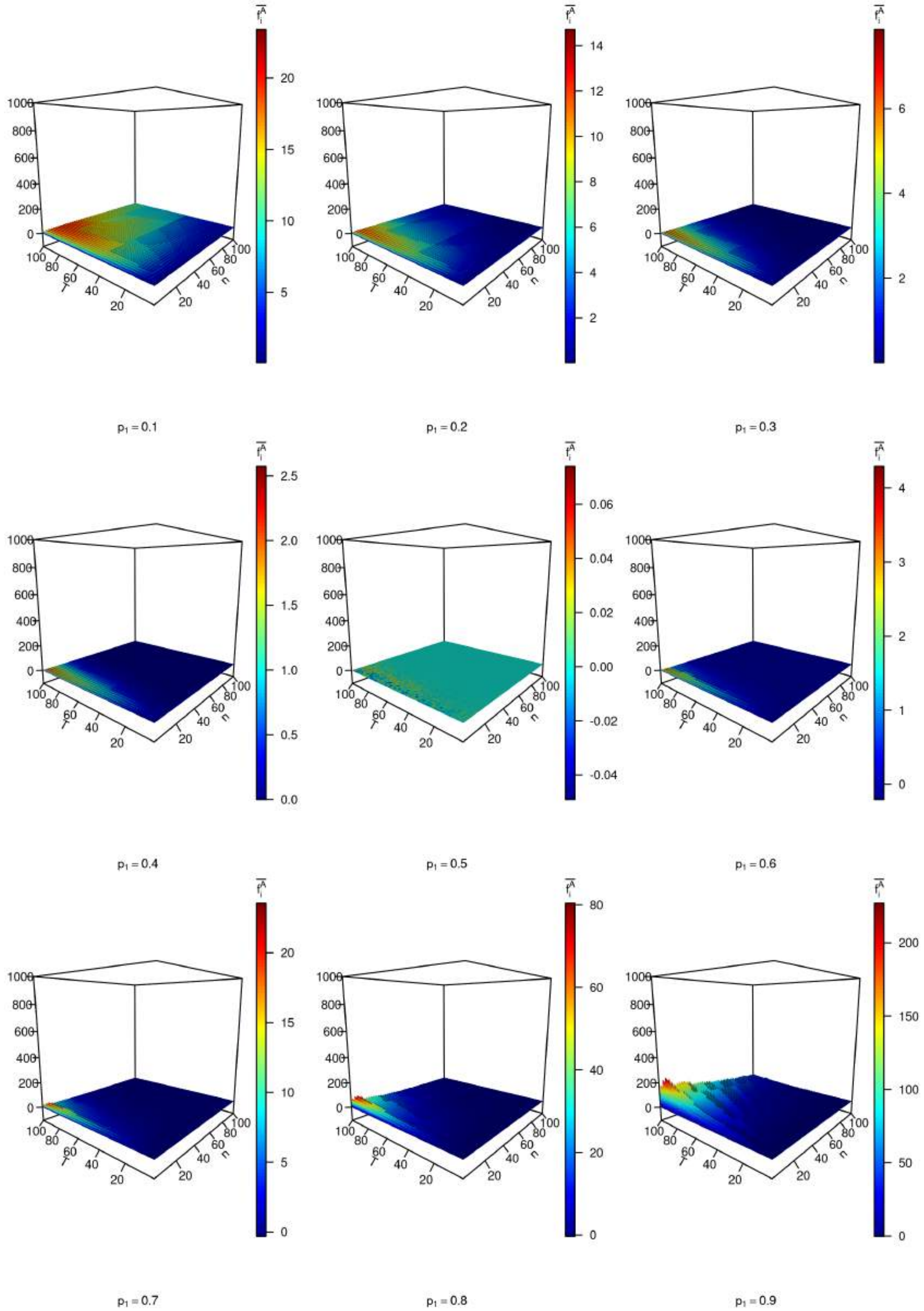

Figure A.5: Trading route 1.1, $w = 0.6$, noise $p_2 = 0.5$.

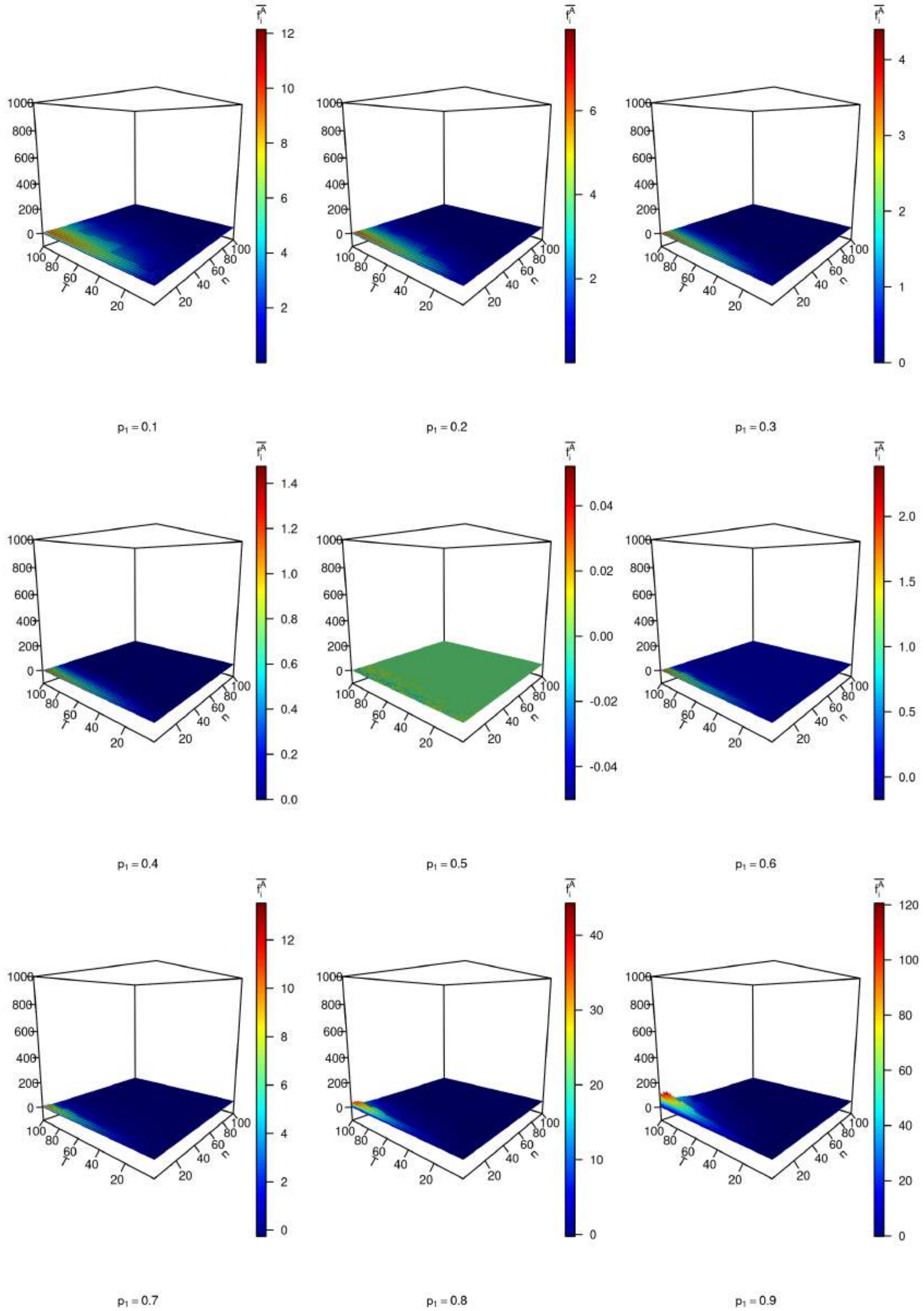

Figure A.6: Trading route 1.1, $w = 0.6$, noise $p_2 = 0.6$.

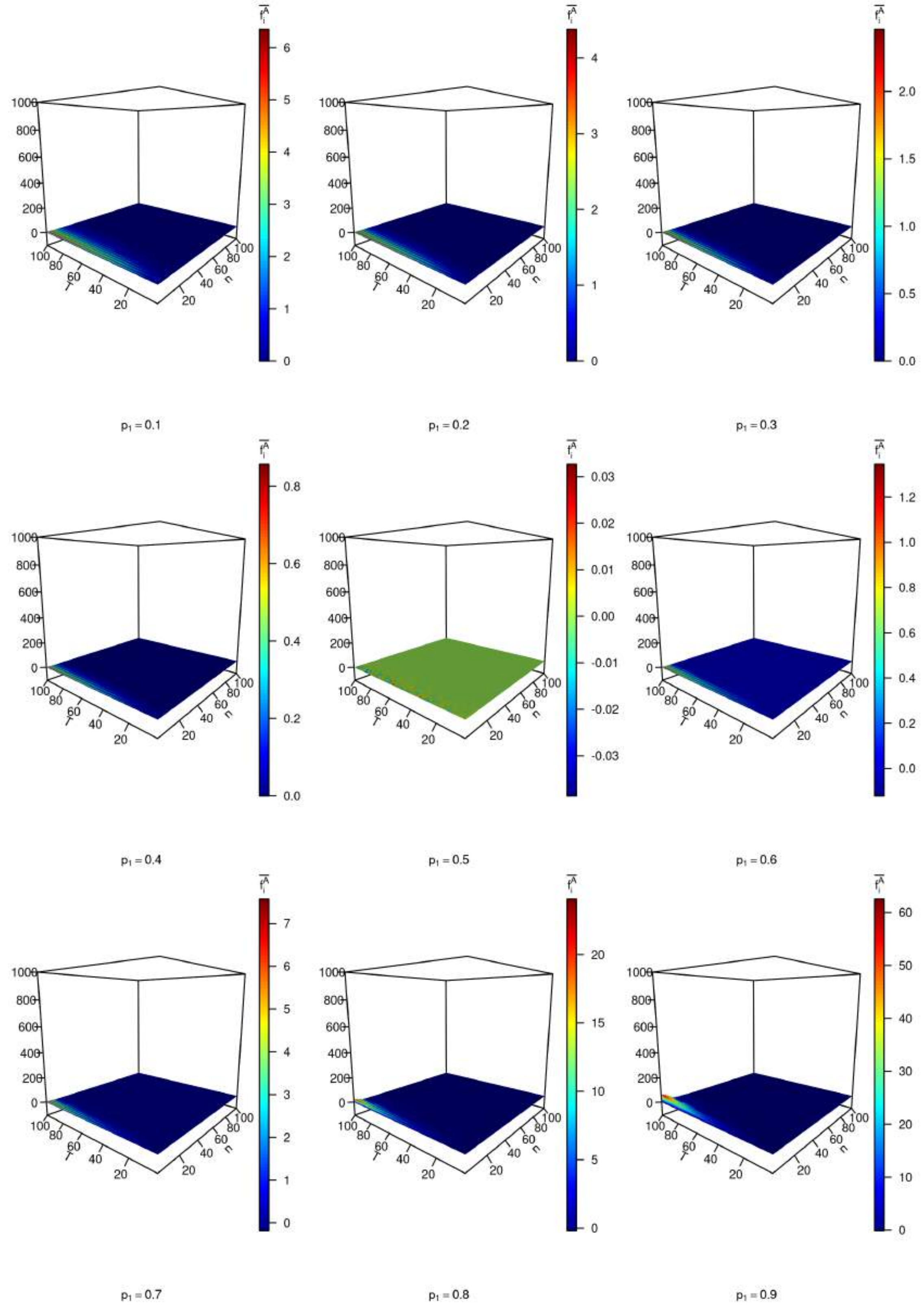

Figure A.7: Trading route 1.1, $w = 0.6$, noise $p_2 = 0.7$.

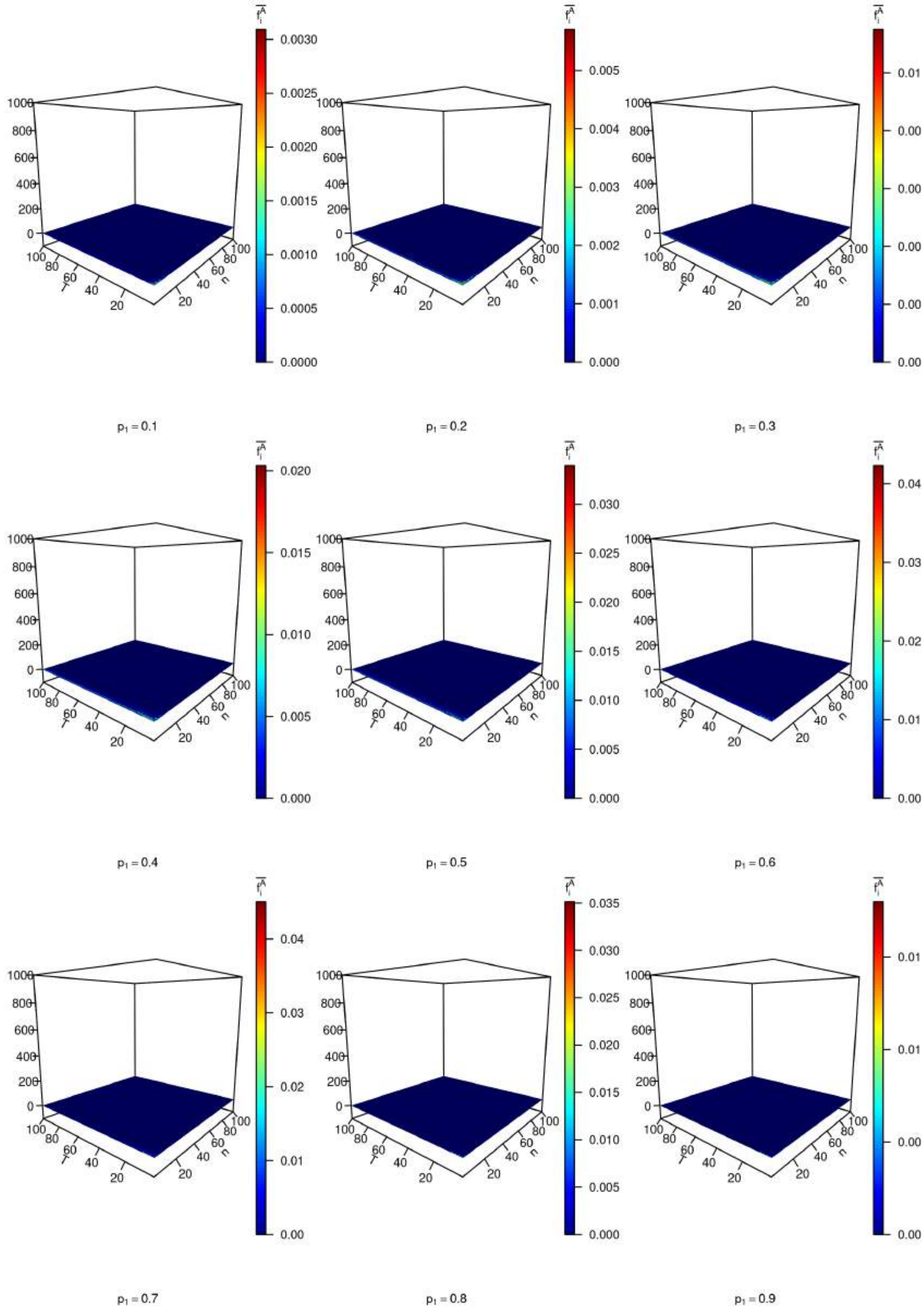

Figure A.8: Trading route 1.1, $w = 0.6$, noise $p_2 = 0.8$.

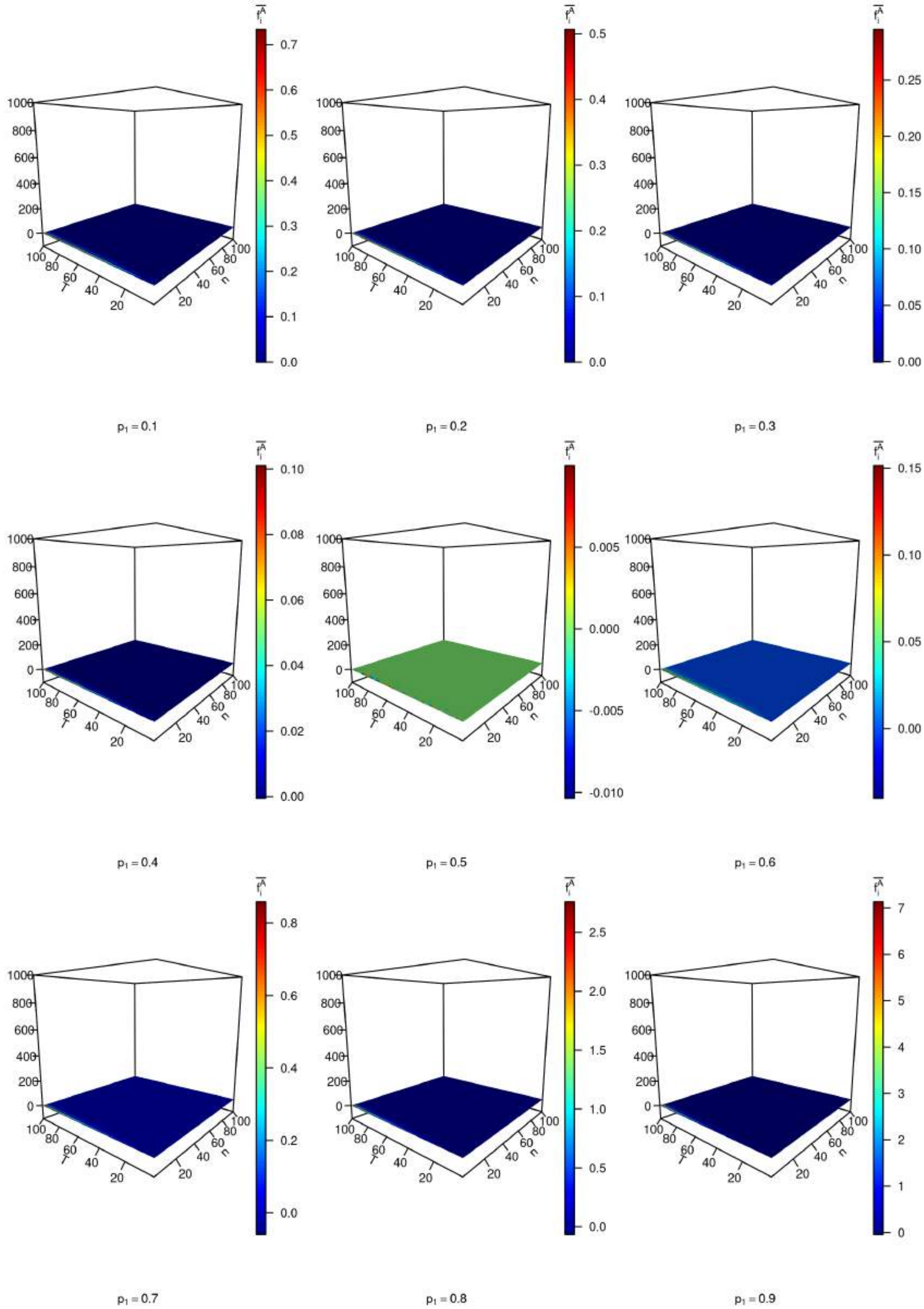

Figure A.9: Trading route 1.1, $w = 0.6$, noise $p_2 = 0.9$.

Appendix B

Performance evaluation of NPI European call option trading routes

This appendix presents a full example of average present value payoff surface $\overline{f_i^C}$ of European call option trading route 2.1 with threshold parameter value $w = 0.6$ under the different market conditions and subject to the different noise levels. The example is part of simulation results in Chapter 5 which shows that the proposed NPI European call option trading routes' noise recognition capability under different market conditions and data learning ability under low level noise affection. s mentions in Chapter 5, trading route 2.1 and 2.2 have similar decaying phenomenons in surface $\overline{f_i^C}$. The presented example could be regarded as the general property demonstrations for both route 2.1 and route 2.2.

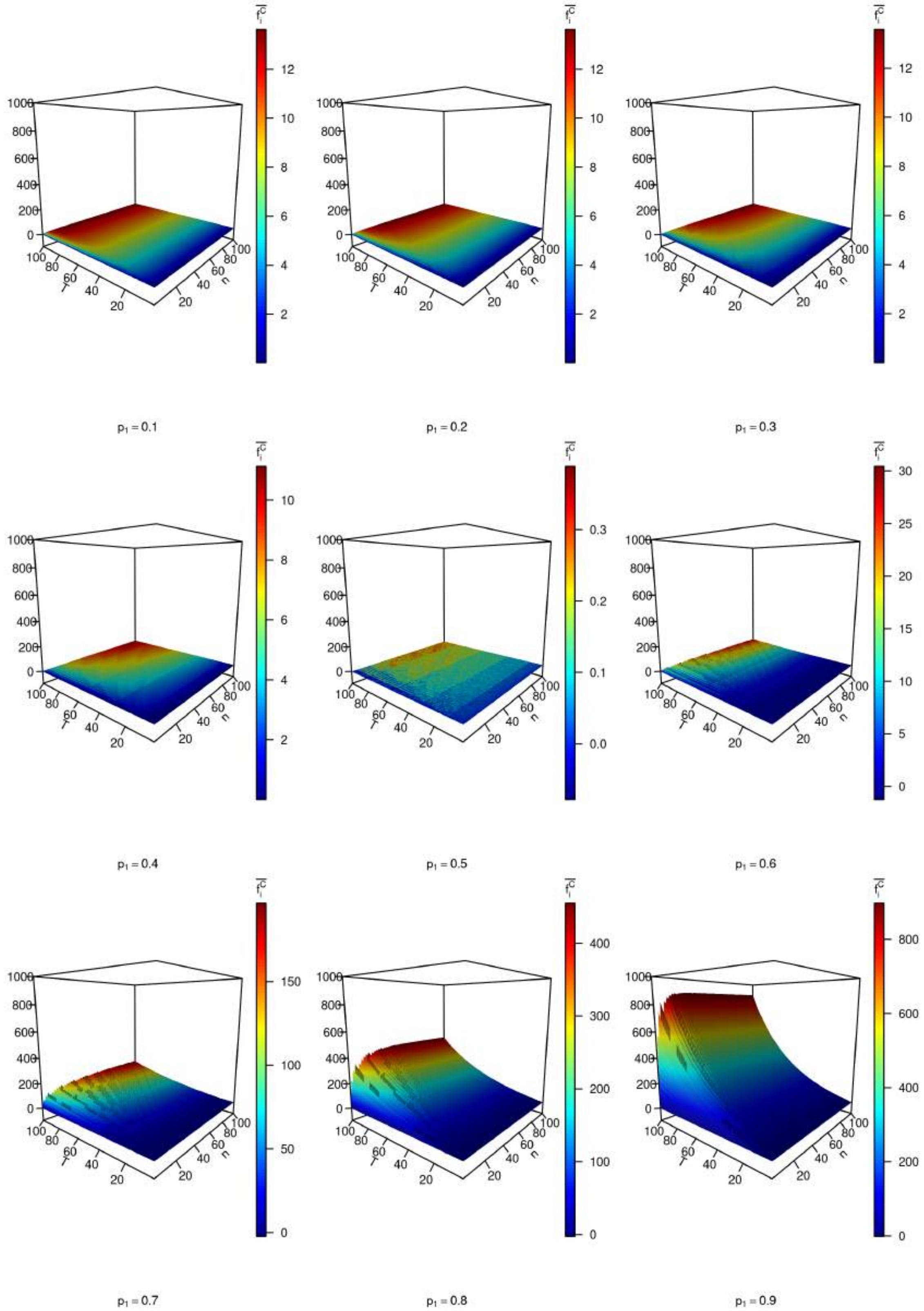

Figure B.1: Trading route 2.1, $w = 0.6$, noise $p_2 = 0.1$.

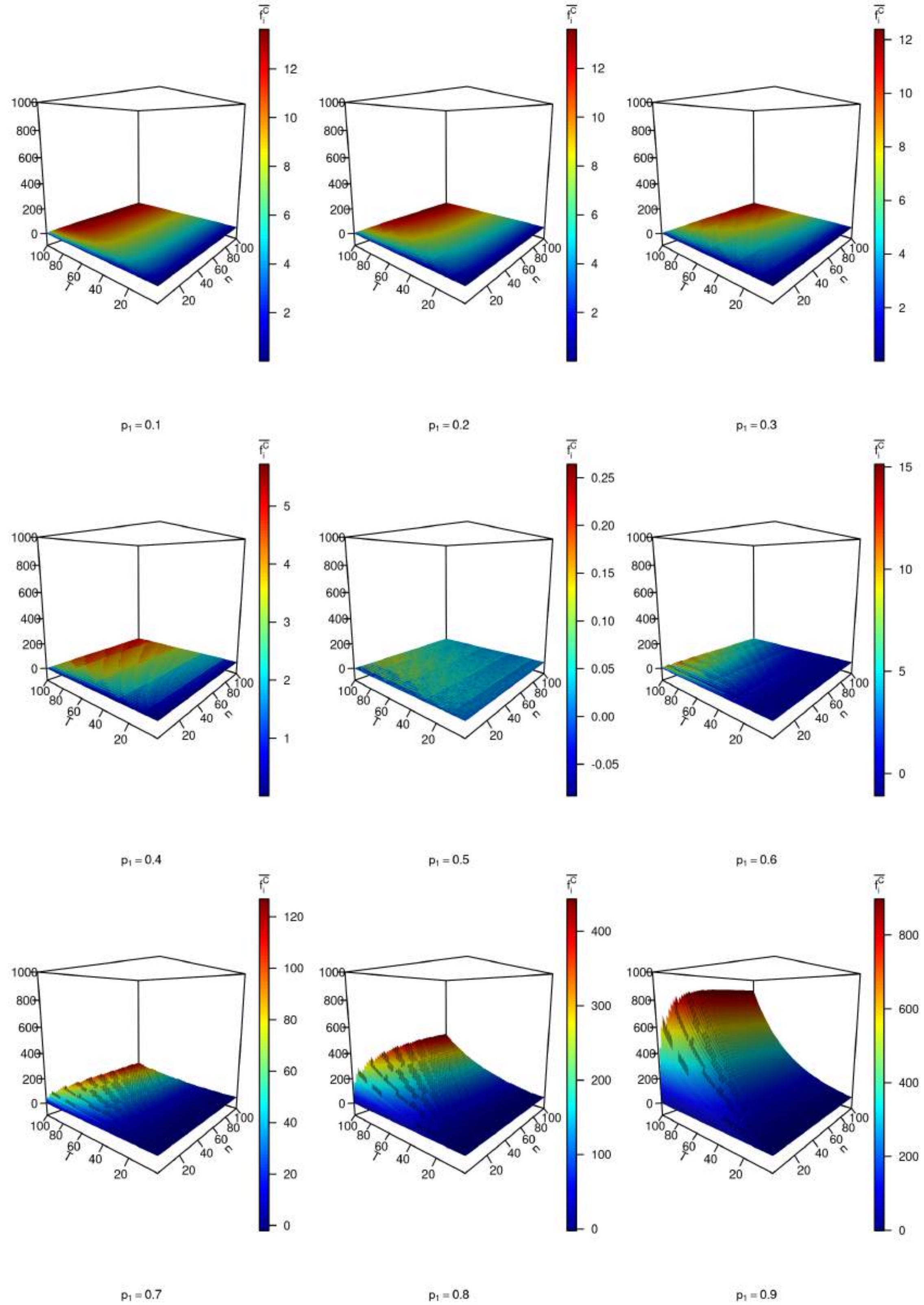

Figure B.2: Trading route 2.1, $w = 0.6$, noise $p_2 = 0.2$.

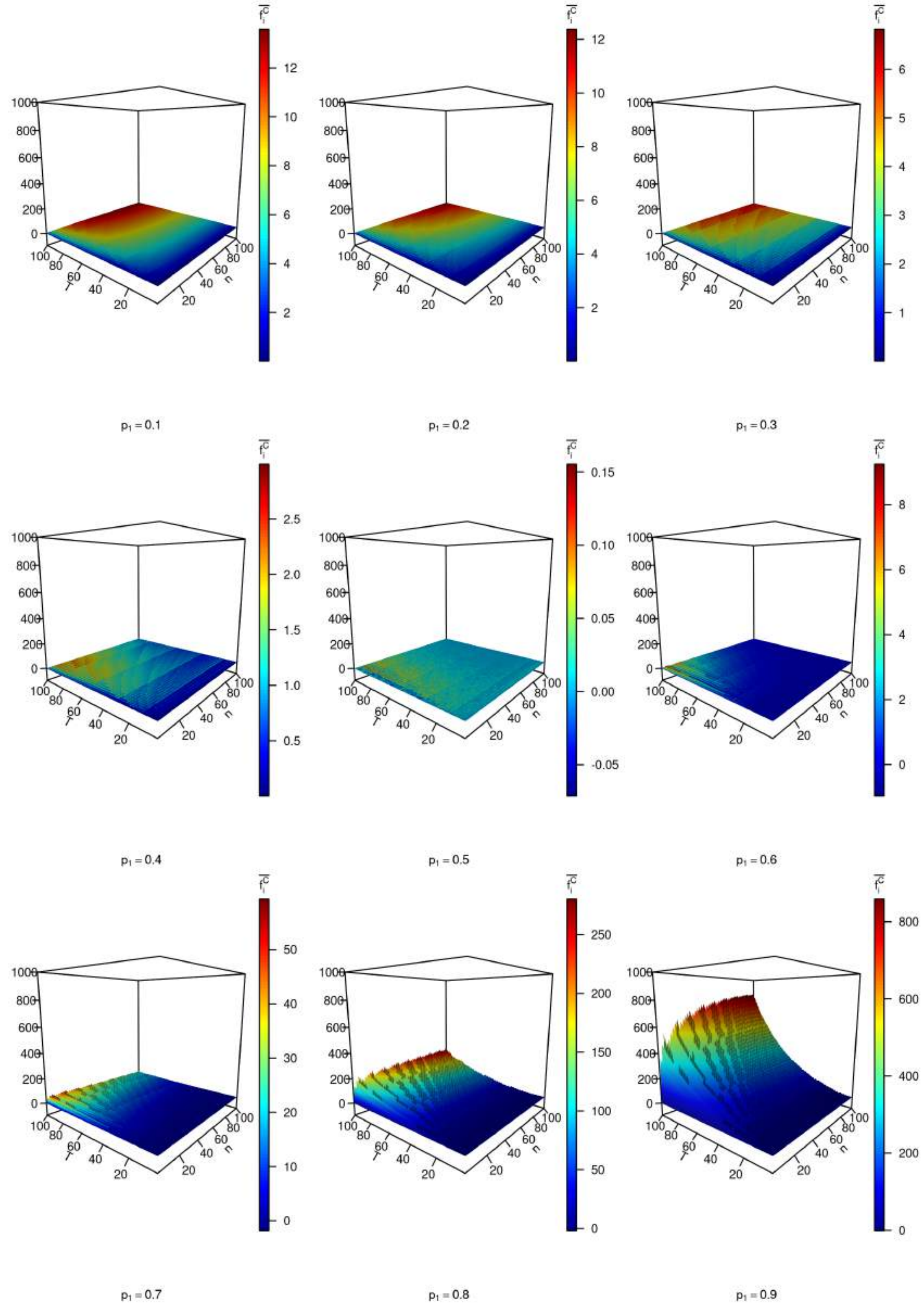

Figure B.3: Trading route 2.1, $w = 0.6$, noise $p_2 = 0.3$.

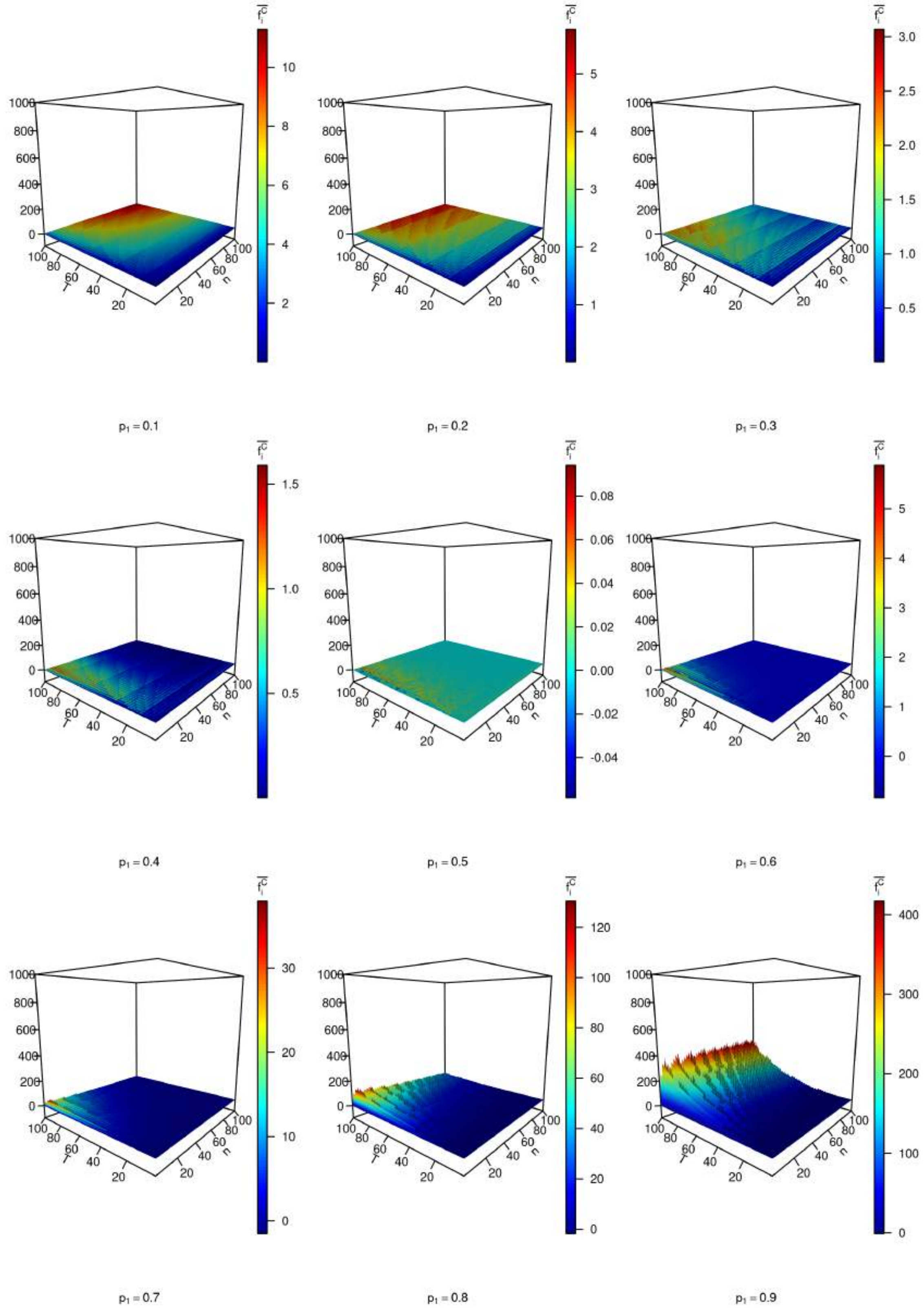

Figure B.4: Trading route 2.1, $w = 0.6$, noise $p_2 = 0.4$.

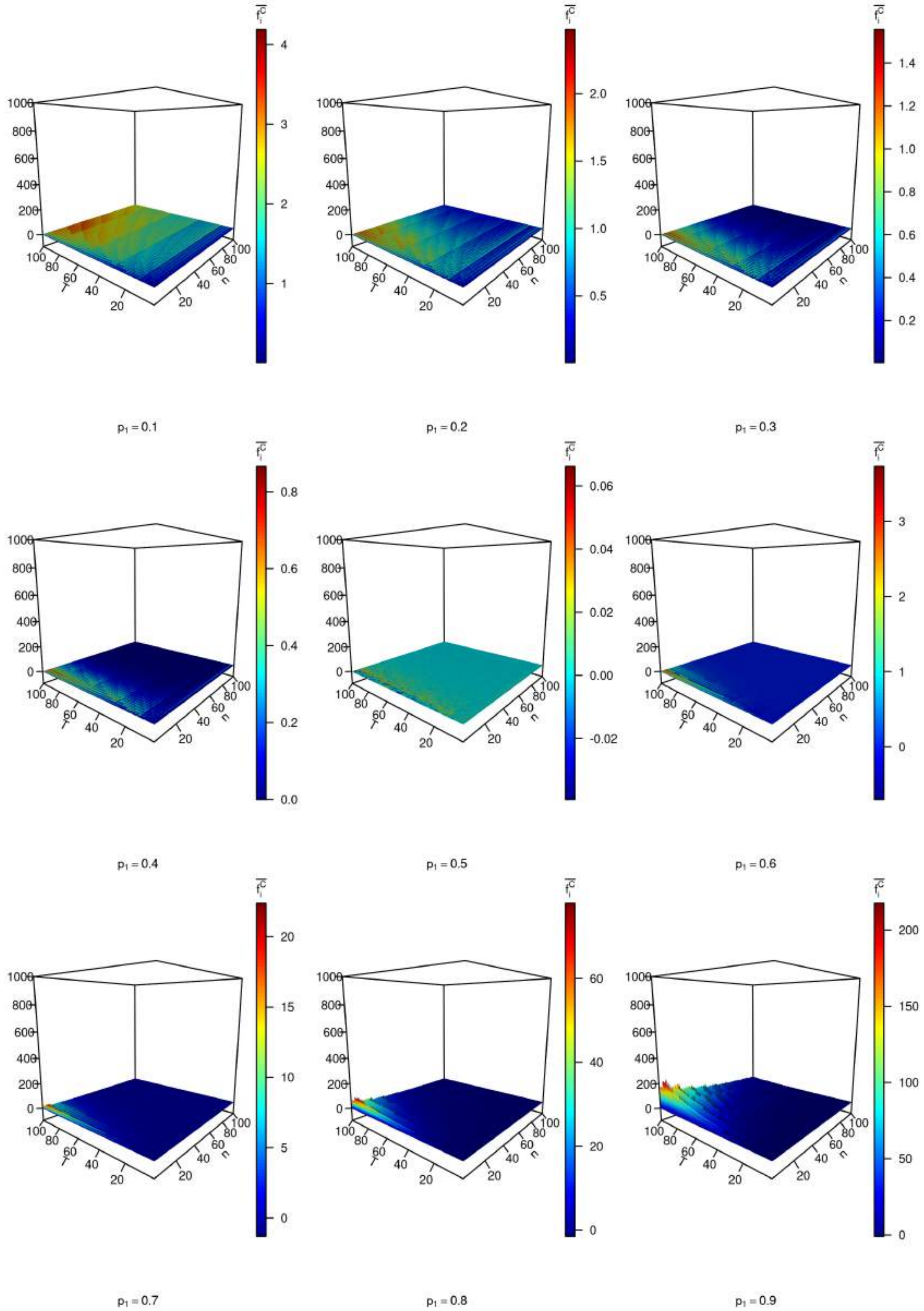

Figure B.5: Trading route 2.1, $w = 0.6$, noise $p_2 = 0.5$.

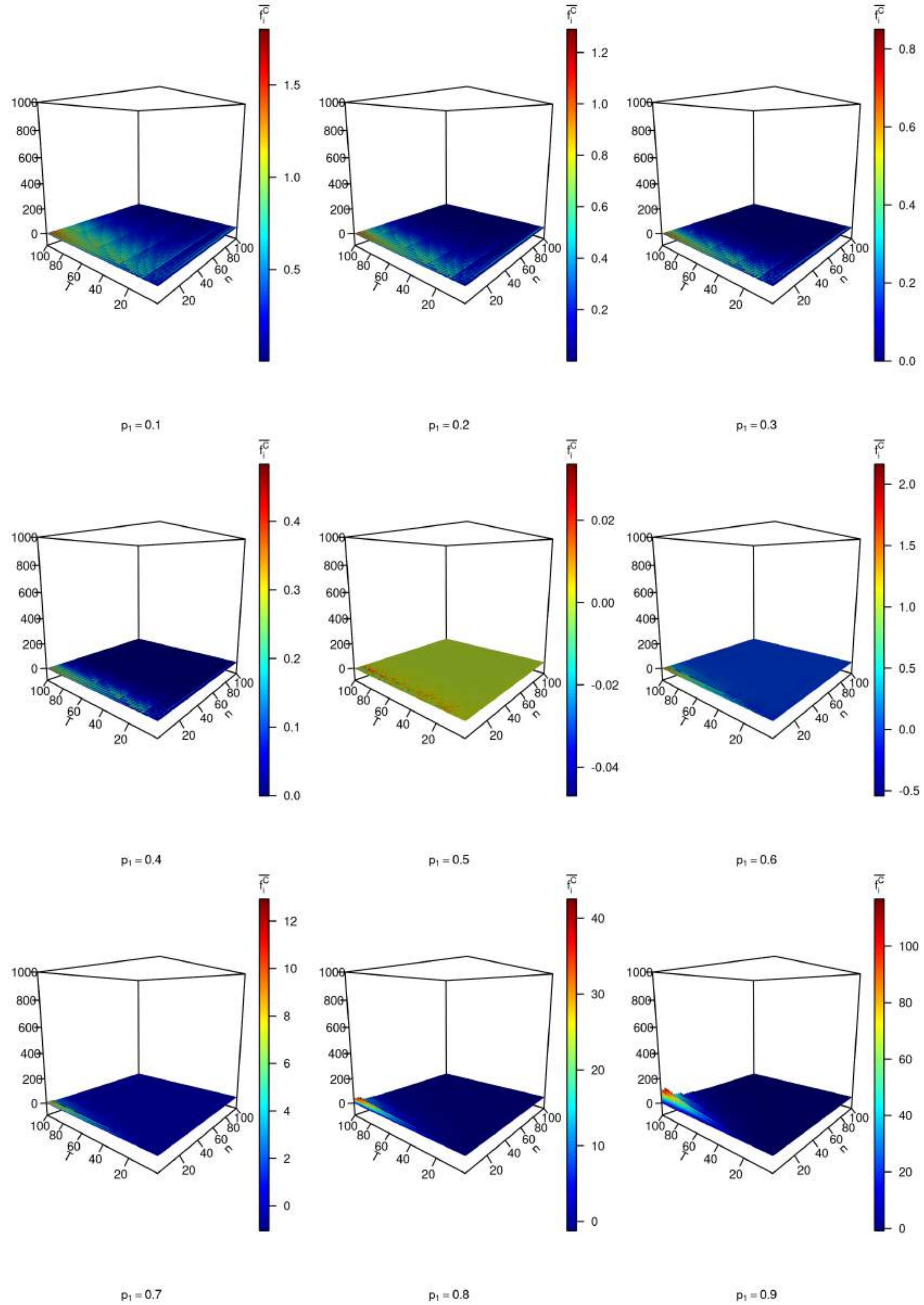

Figure B.6: Trading route 2.1, $w = 0.6$, noise $p_2 = 0.6$.

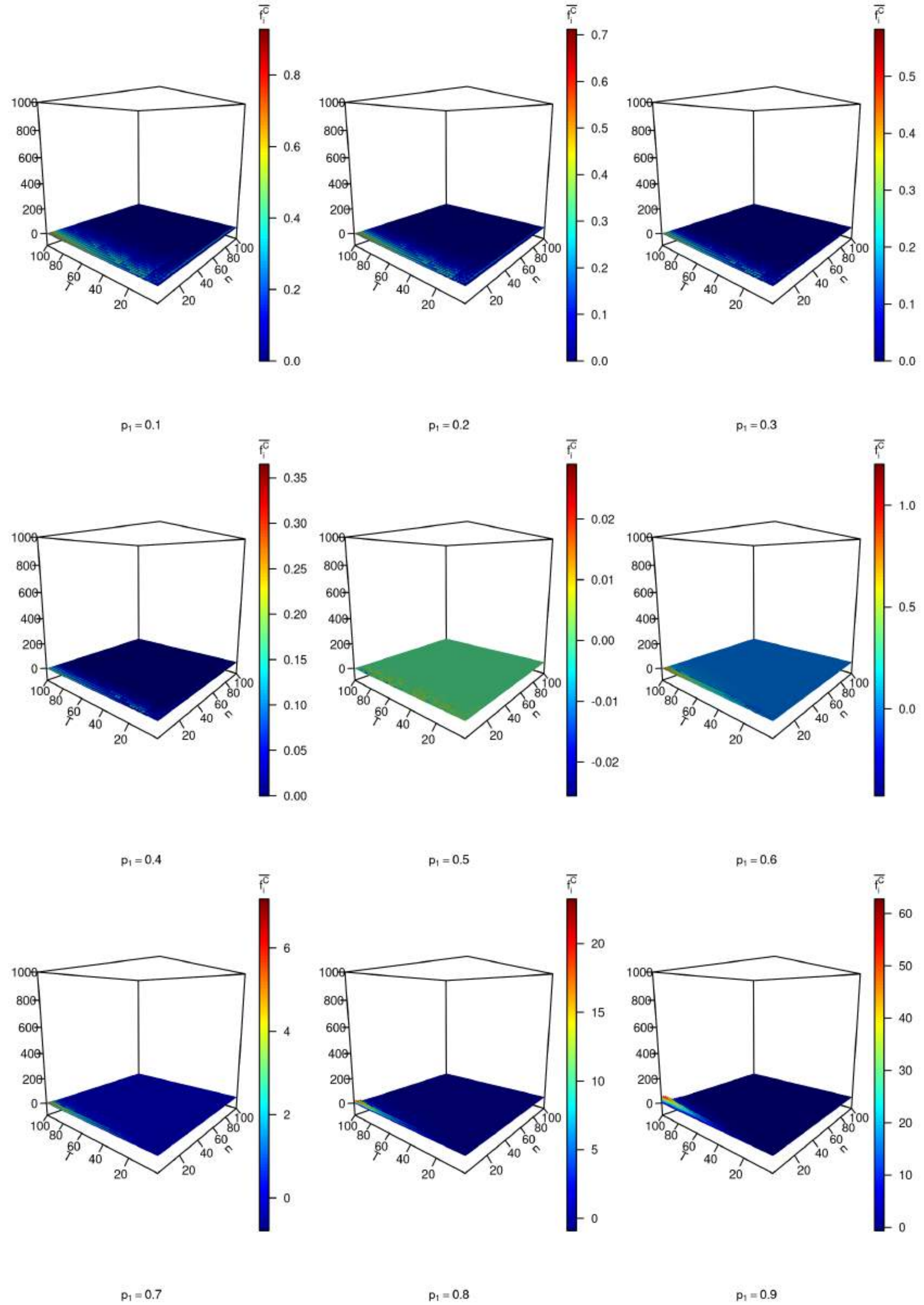

Figure B.7: Trading route 2.1, $w = 0.6$, noise $p_2 = 0.7$.

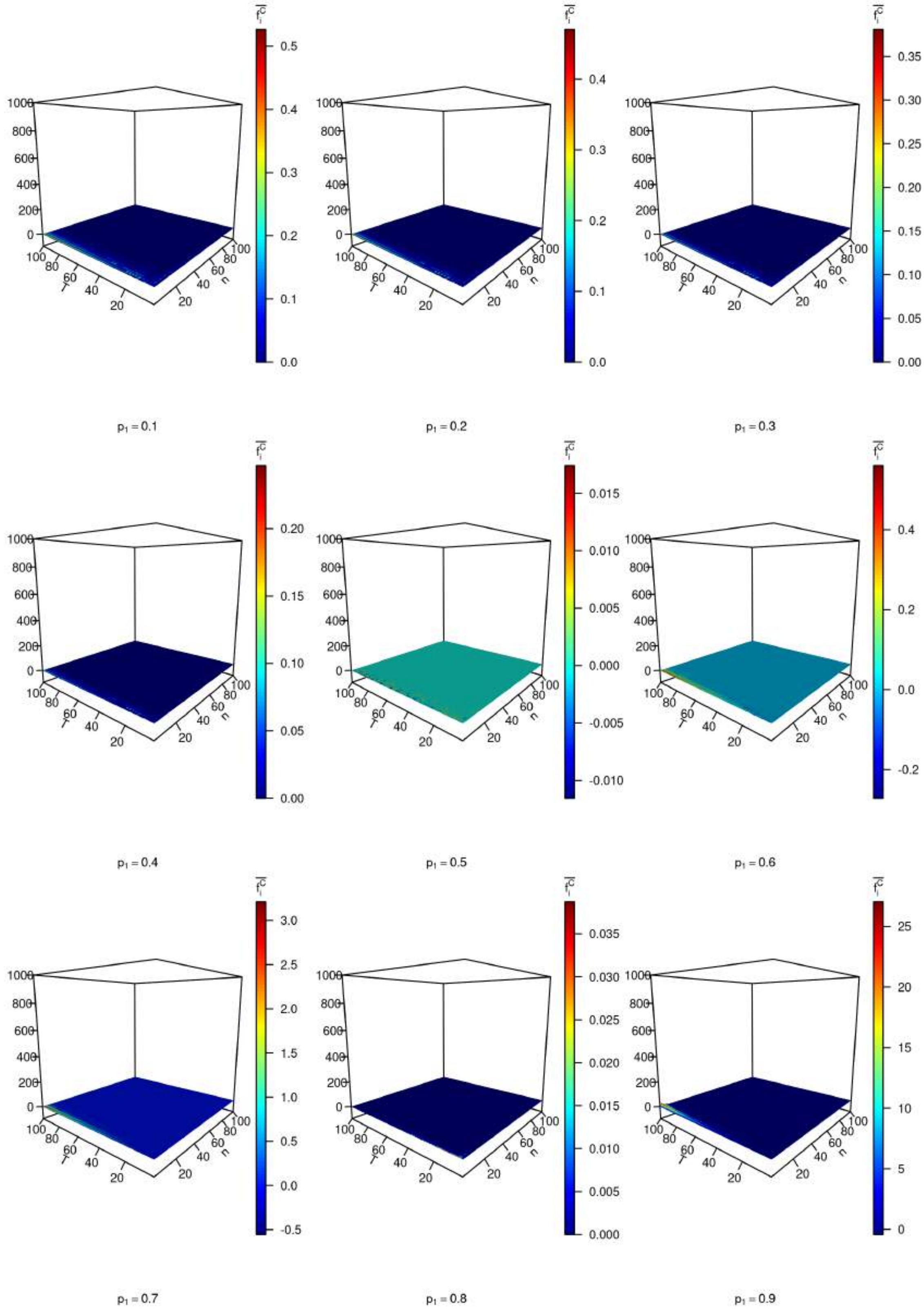

Figure B.8: Trading route 2.1, $w = 0.6$, noise $p_2 = 0.8$.

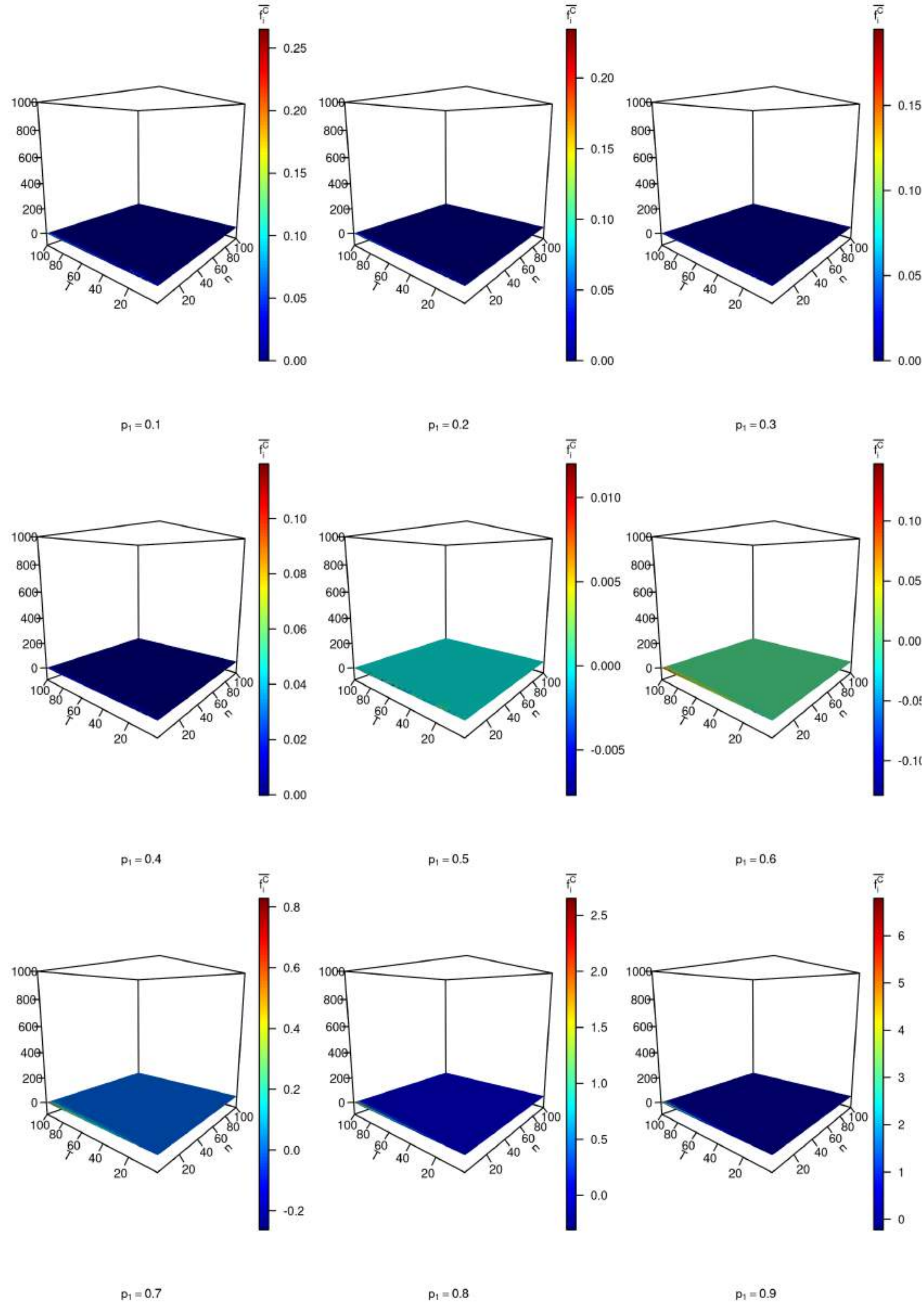

Figure B.9: Trading route 2.1, $w = 0.6$, noise $p_2 = 0.9$.

Appendix C

Performance evaluation of NPI European put option trading routes

This appendix presents a full example of average present value payoff surface $\overline{f_i^P}$ of European put option trading route 2.3 with threshold parameter value $w = 0.6$ under the different market conditions and subject to the different noise levels. The example is part of simulation results in Chapter 5 which shows that the proposed NPI European put option trading routes' noise recognition capability under different market conditions and data learning ability under low level noise affection. As mentions in Chapter 5, trading route 2.3 and 2.4 have similar decaying phenomenons in surface $\overline{f_i^P}$. The presented example could be regarded as the general property demonstrations for both route 2.3 and route 2.4.

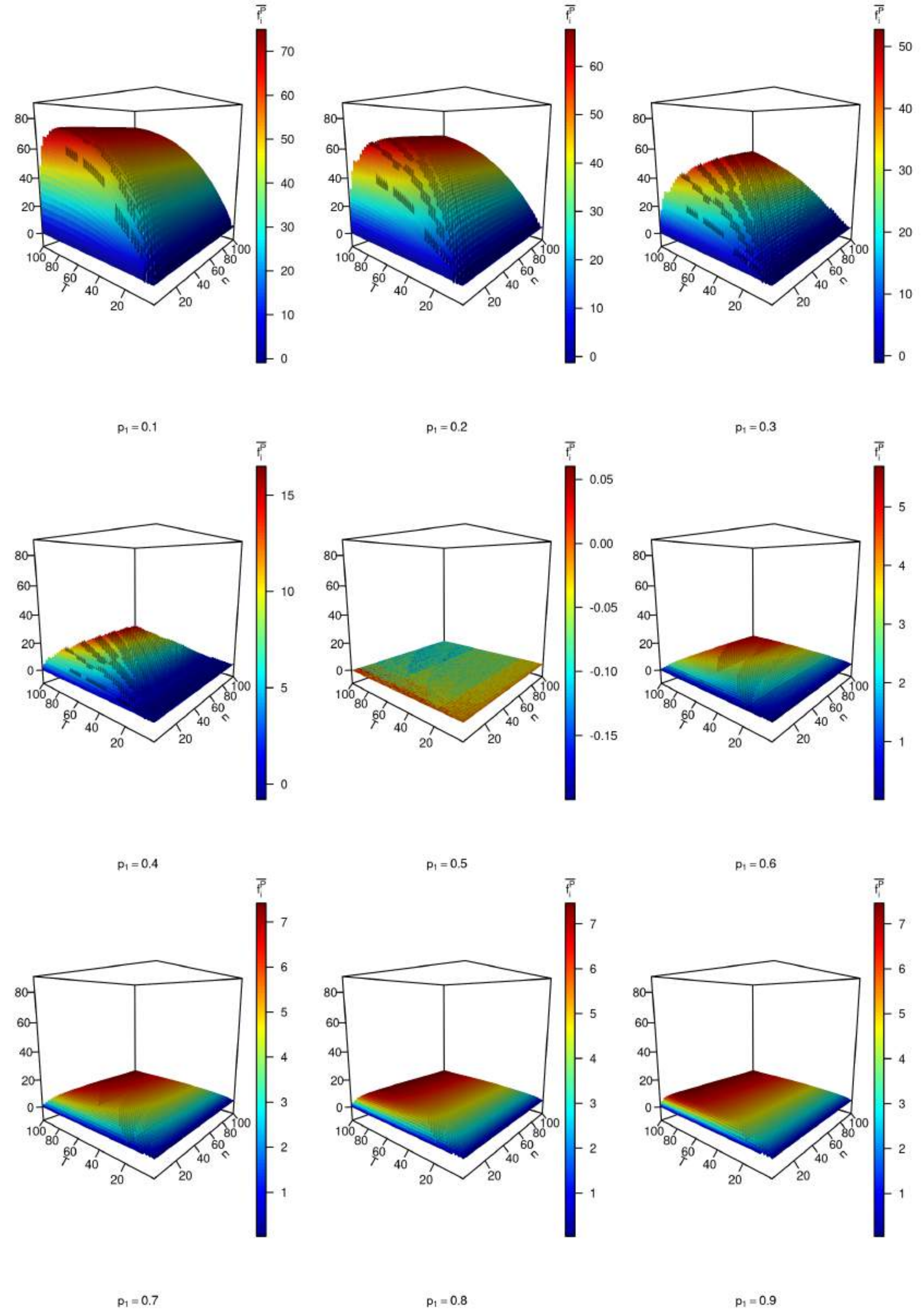

Figure C.1: Trading route 2.3, $w = 0.6$, noise $p_2 = 0.1$.

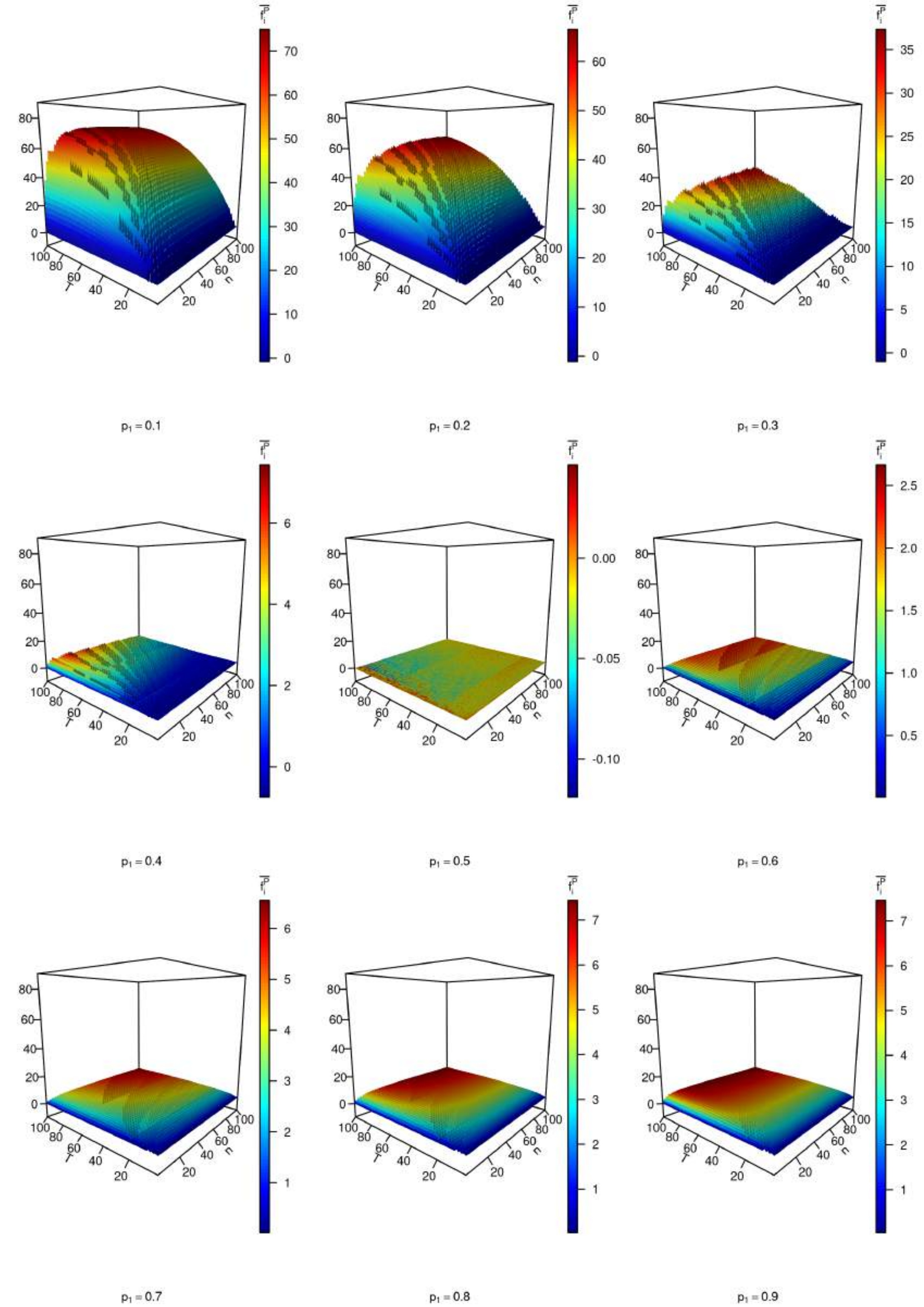

Figure C.2: Trading route 2.3, $w = 0.6$, noise $p_2 = 0.2$.

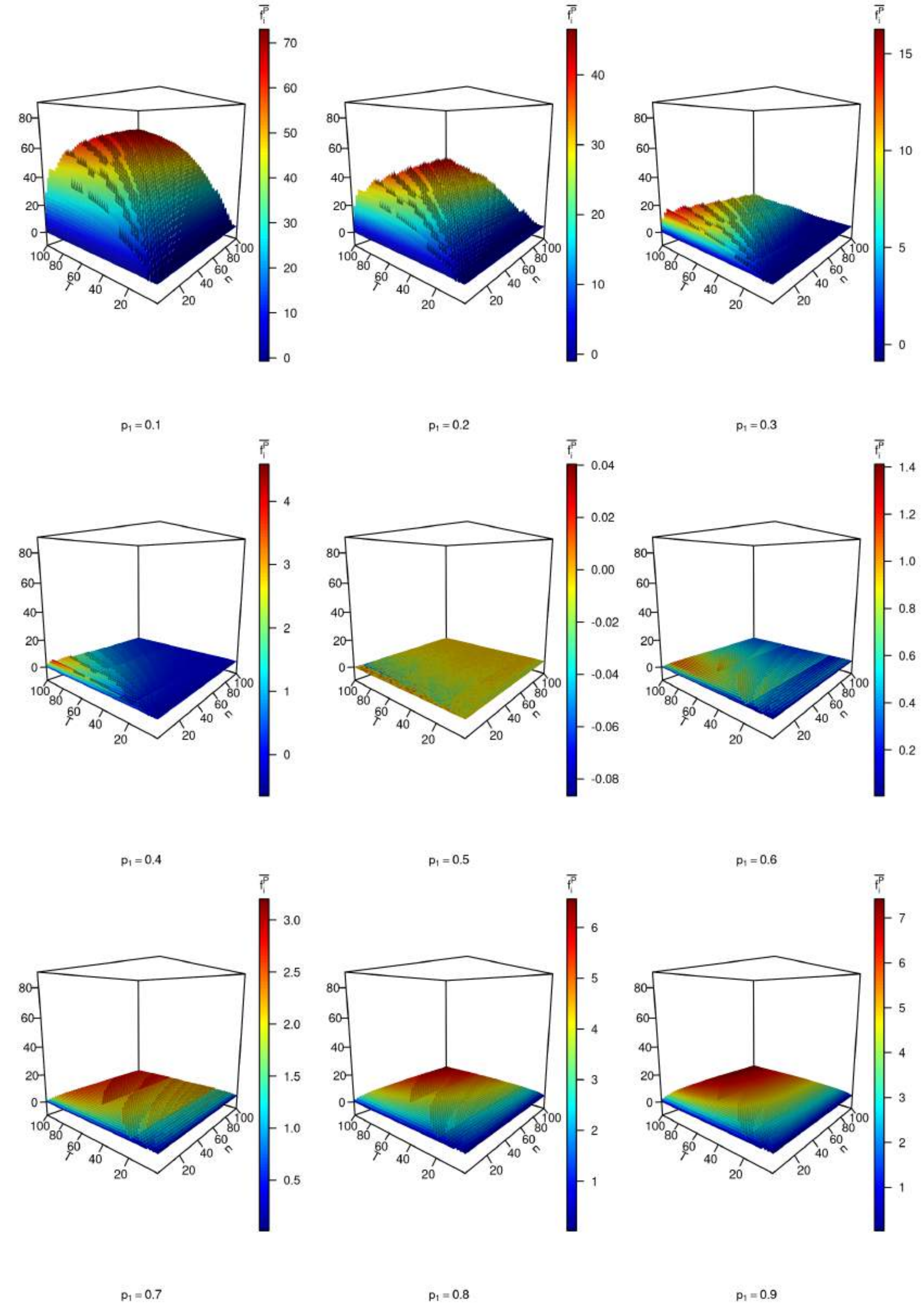

Figure C.3: Trading route 2.3, $w = 0.6$, noise $p_2 = 0.3$.

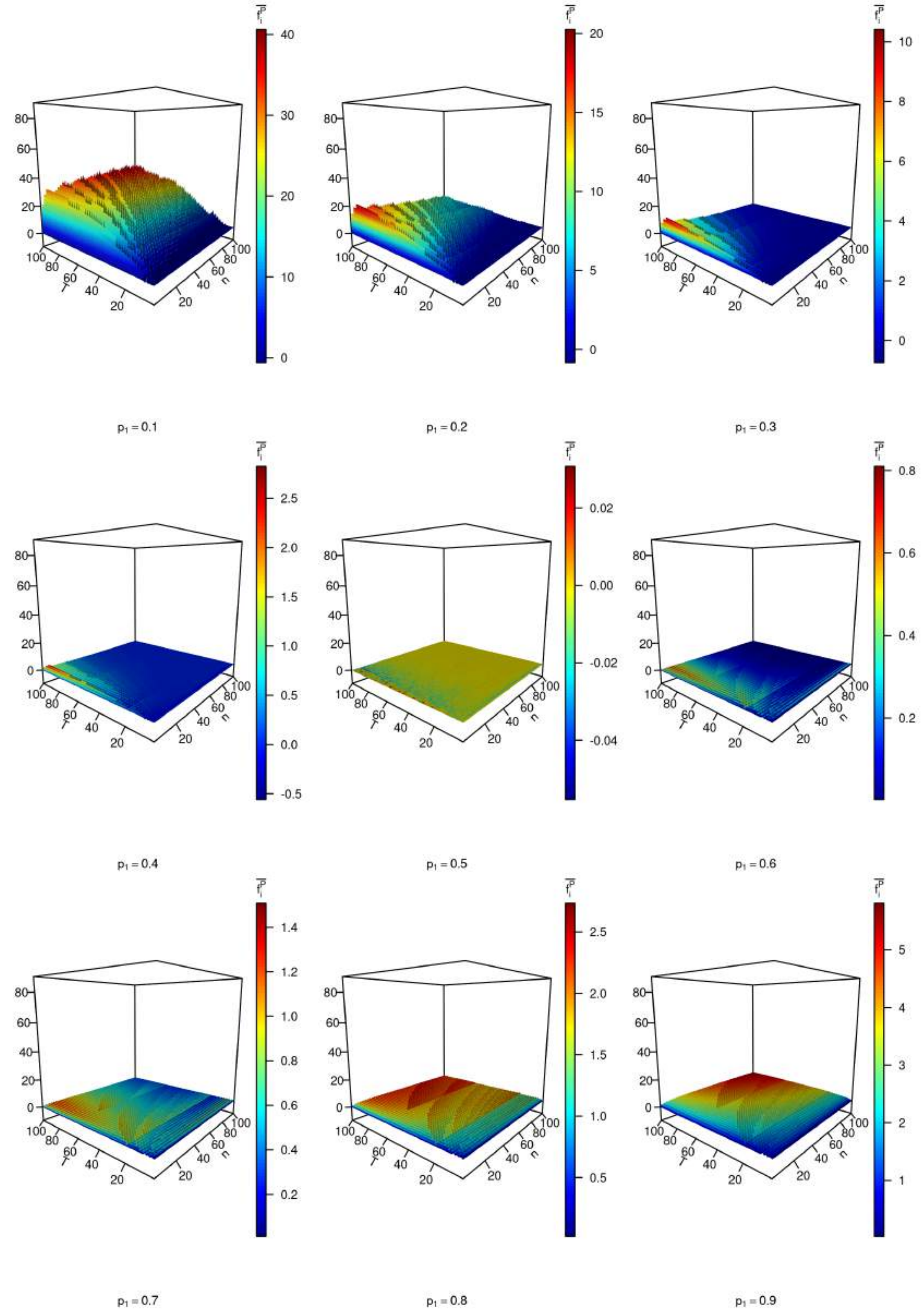

Figure C.4: Trading route 2.3, $w = 0.6$, noise $p_2 = 0.4$.

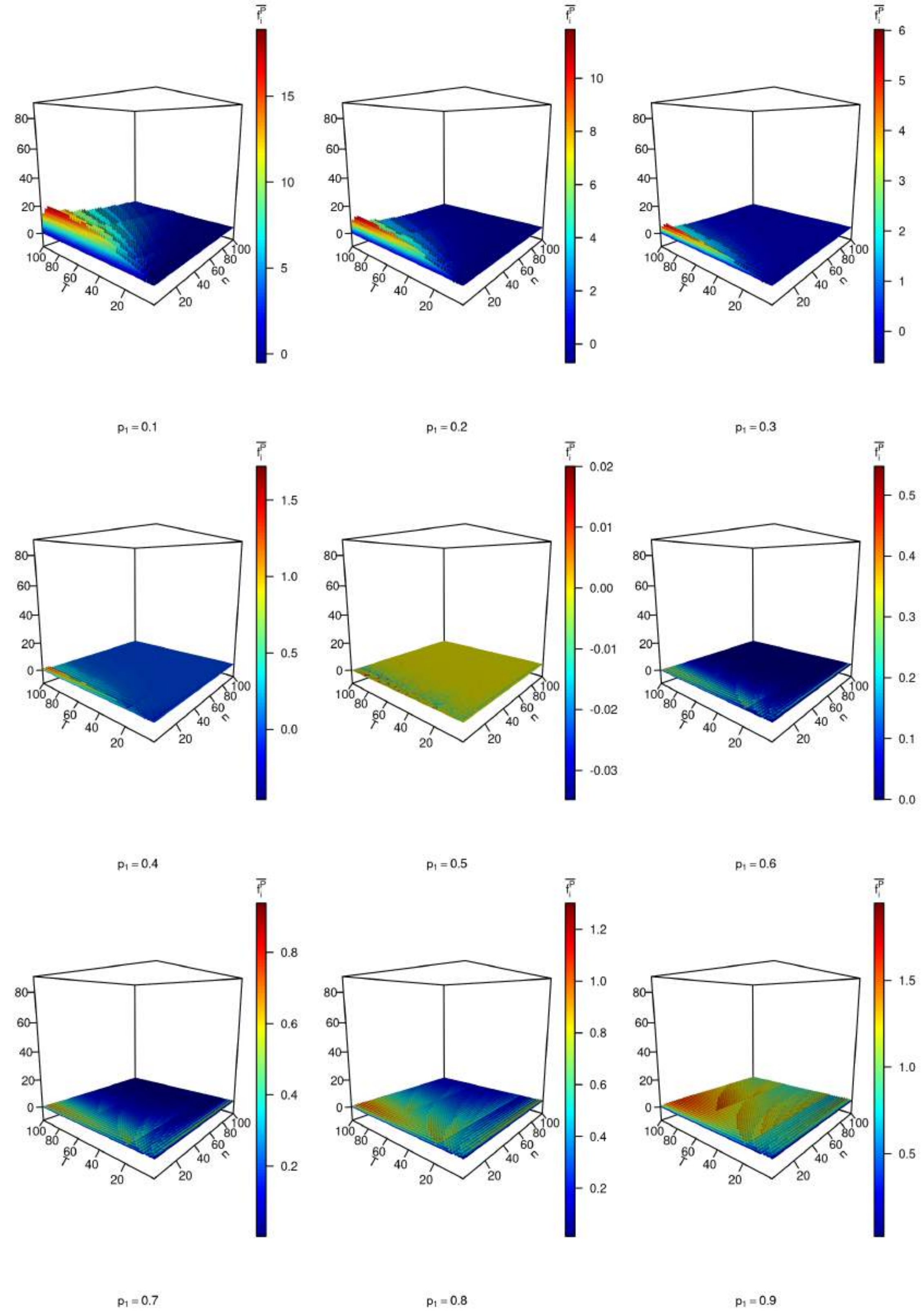

Figure C.5: Trading route 2.3, $w = 0.6$, noise $p_2 = 0.5$.

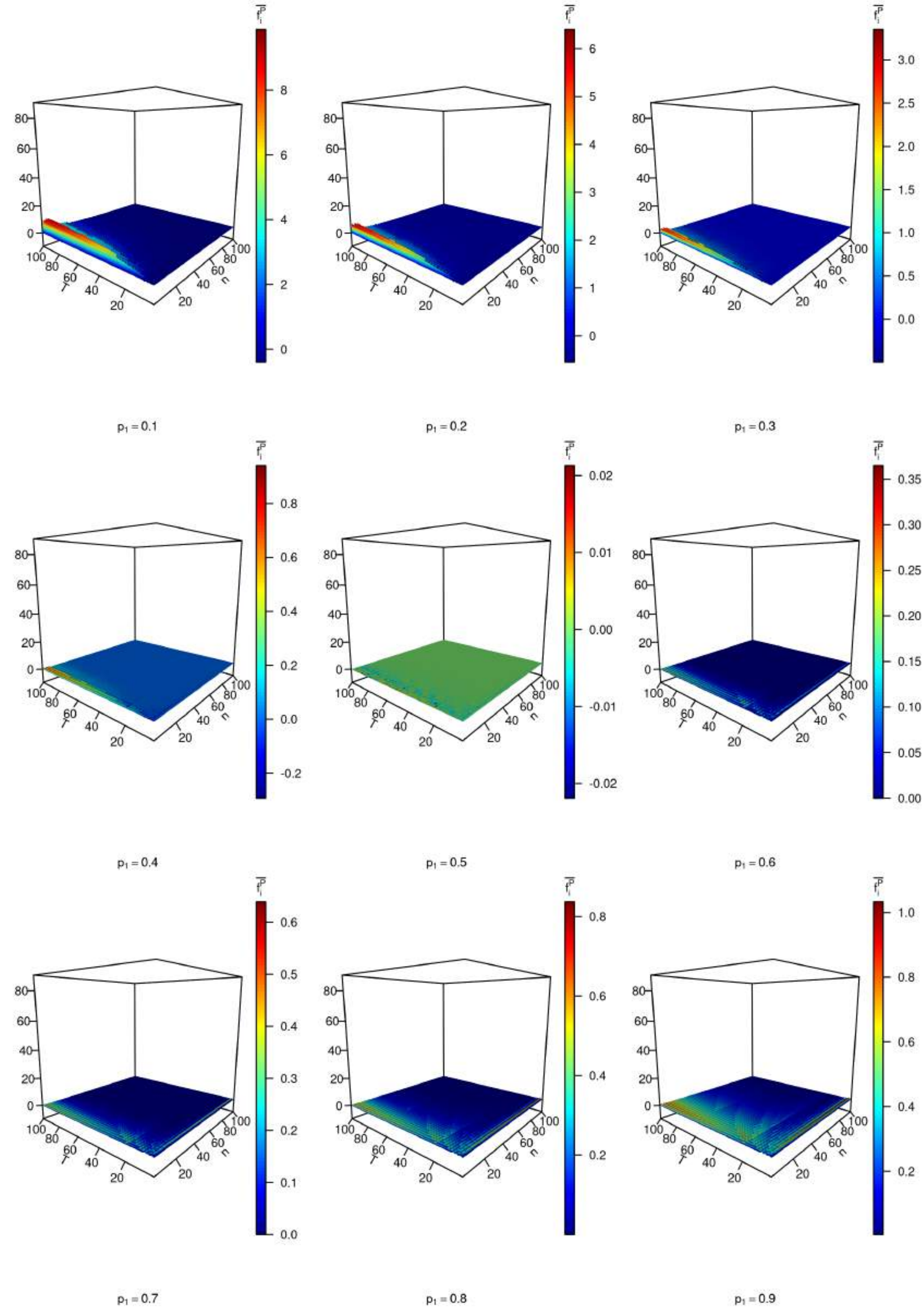

Figure C.6: Trading route 2.3, $w = 0.6$, noise $p_2 = 0.6$.

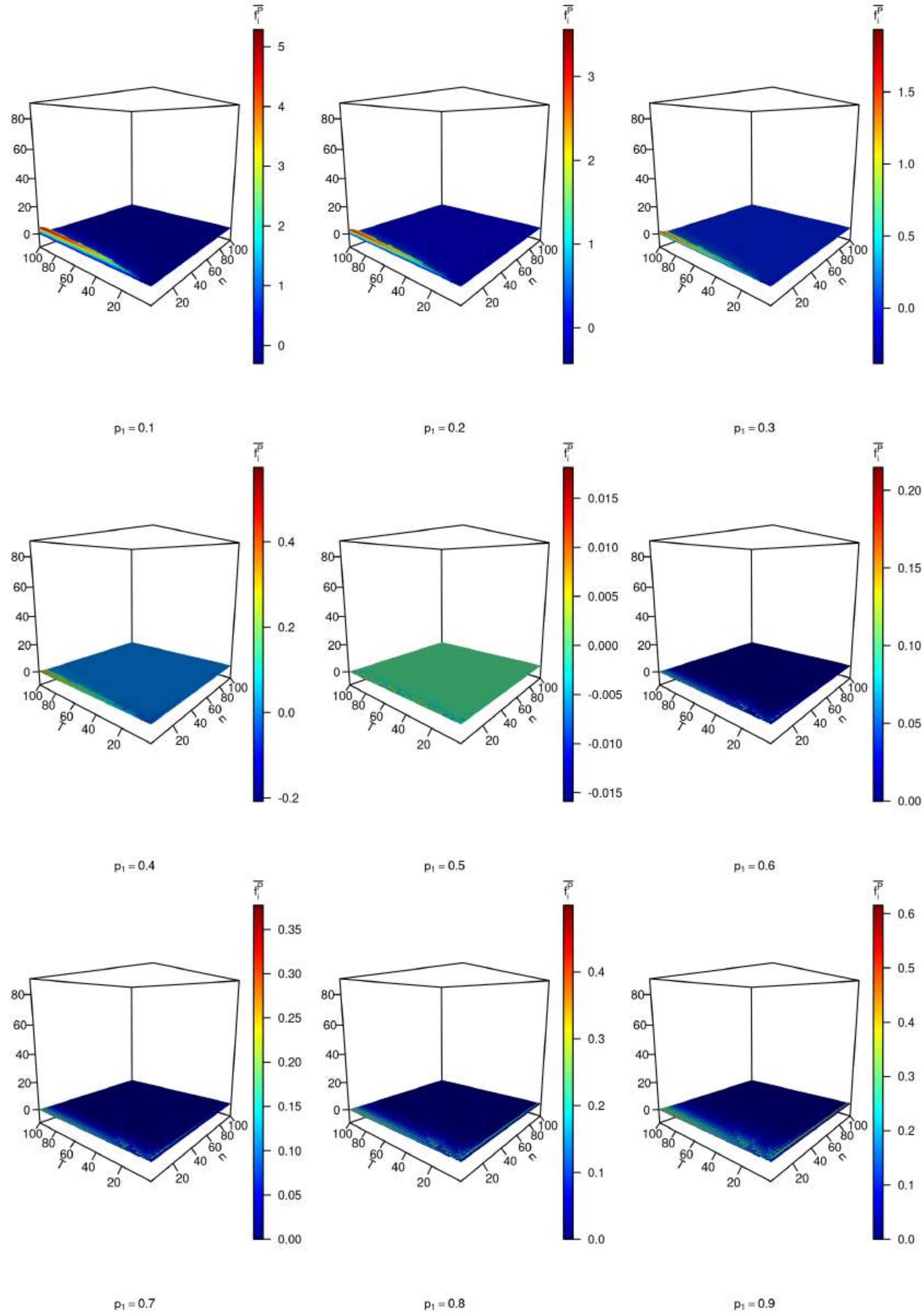

Figure C.7: Trading route 2.3, $w = 0.6$, noise $p_2 = 0.7$.

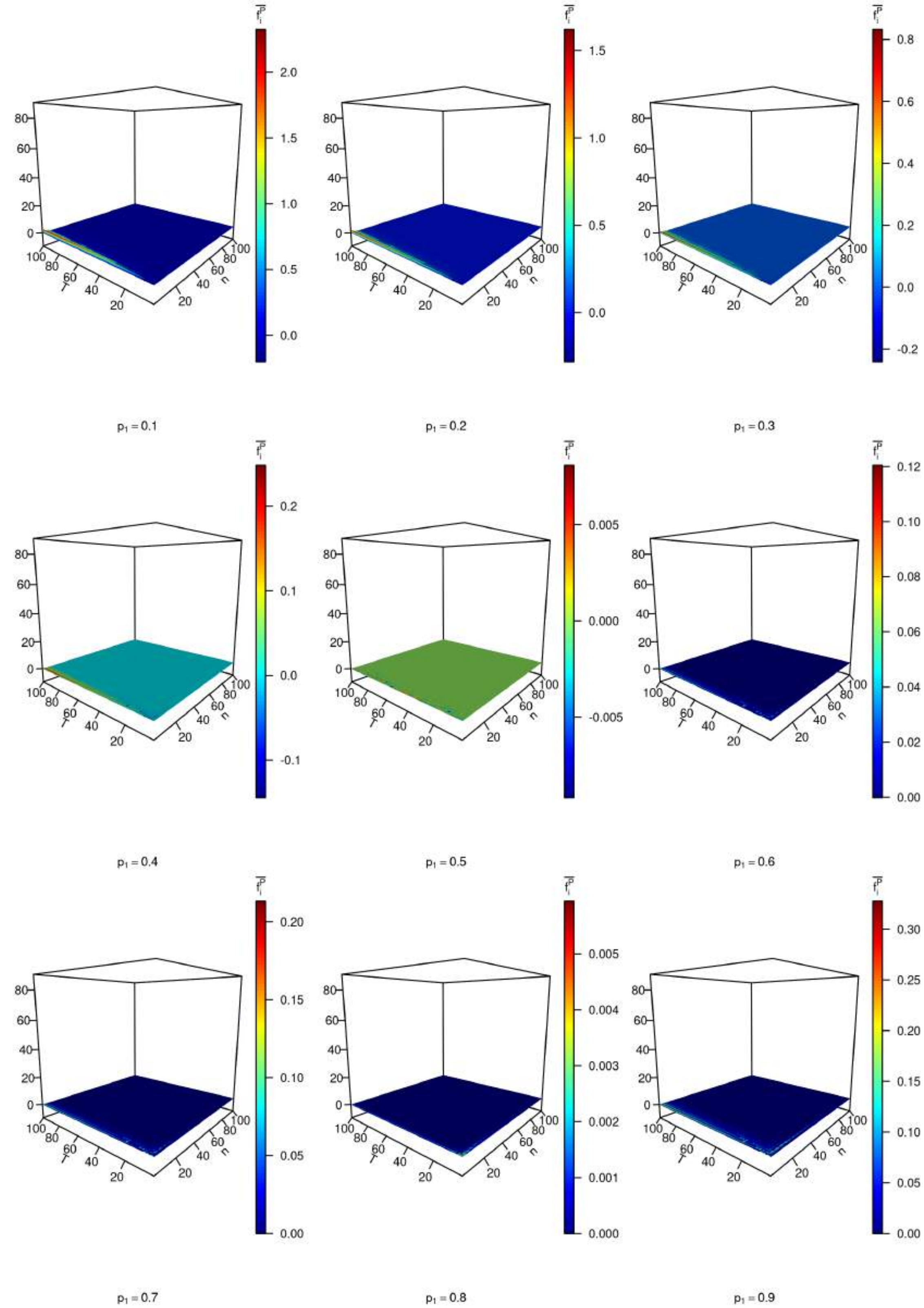

Figure C.8: Trading route 2.3, $w = 0.6$, noise $p_2 = 0.8$.

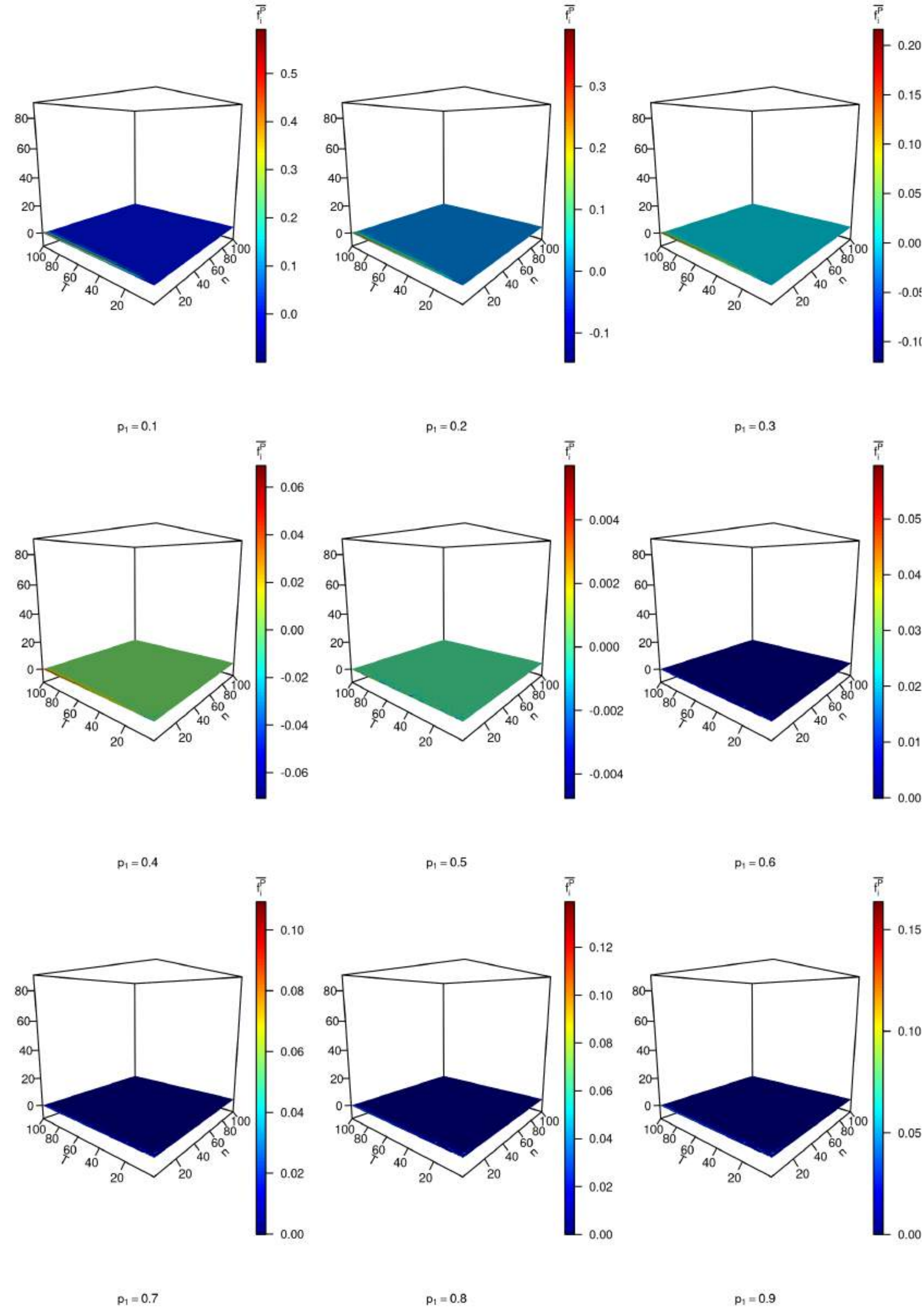

Figure C.9: Trading route 2.3, $w = 0.6$, noise $p_2 = 0.9$.